

Integrative Analyse des cyanobakteriellen Stoffwechsels

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

im Fach Biologie

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät I

der Humboldt-Universität zu Berlin

von

M.Sc. Henning Knoop



Präsident der Humboldt-Universität zu Berlin

Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I

Prof. Stefan Hecht, Ph.D.

Gutachter:

1. Prof. Dr. Wolfgang Lockau
2. Prof. Dr. Martin Hagemann
3. Dr. Ralf Steuer

Tag der mündlichen Prüfung: 16. Oktober 2014

Zusammenfassung

Cyanobakterien sind einzellige, phototrophe Mikroorganismen, die in unterschiedlichsten Umgebungen und Ökosystemen vorkommen. Aufgrund ihrer Fähigkeit, Sonnenenergie und atmosphärisches Kohlenstoffdioxid für das Zellwachstum zu nutzen, stellen Cyanobakterien einen idealen Modellorganismus für ein besseres Verständnis der oxygenen Photosynthese und des phototrophen Stoffwechsels dar. Sie werden auch zunehmend als potenzieller Wirtsorganismus für die Synthese von wertvollen Chemikalien und verschiedenen Biokraftstoffen wahrgenommen. Um diese Vorteile in vollem Umfang zu nutzen, bietet die genomskalige Rekonstruktion von Mikroorganismen ein weitgehendes Verständnis und eine Wissensbasis über die metabolischen Umwandlungen, die während des phototrophen Wachstums stattfinden.

Im Verlauf dieser Arbeit wurde eine detaillierte Rekonstruktion und Analyse des metabolischen Netzwerkes der einzelligen Cyanobakterien *Synechocystis* sp. PCC 6803 und *Synechococcus elongatus* PCC 7942 erstellt, miteinander verglichen und Gemeinsamkeiten und Unterschiede diskutiert. Die Rekonstruktionen basieren dabei auf unterschiedlichen Datenquellen, mehreren neueren Rekonstruktionen des cyanobakteriellen Stoffwechsels und einer ausführlichen manuellen Kuration. Darüber hinaus wurden beim Netzwerk von *Synechocystis* sp. PCC 6803 unklare Reaktionsschritte experimentell validiert und die funktionellen Konsequenzen von unbekannten oder abweichenden Stoffwechselltopologien untersucht. Dabei umfasst das Modell neuartige Ergebnisse in Bezug auf den cyanobakteriellen TCA-Zyklus, einen angeblich vorhandenen Glyoxylatzzyklus und die Rolle der Photorespiration beim Zellwachstum, die mithilfe der Flussbilanzanalyse (*Flux Balance Analysis*) (FBA) systematisch auch in Bezug auf den täglichen Hell-Dunkel-Zyklus des cyanobakteriellen Stoffwechsels gewonnen wurden. Zudem helfen die aus der genomskaligen Rekonstruktion der beiden Organismen gewonnenen Erkenntnisse, Unstimmigkeiten in der derzeitigen Annotation auszumachen.

Ein weiterer Aspekt dieser Arbeit liegt im Zugewinn des allgemeinen Verständnisses über die genomische Diversität innerhalb unterschiedlicher Cyanobakterienstämme und speziell innerhalb des cyanobakteriellen Stoffwechsels. Dazu wurden die Genome mehrerer phototropher Cyanobakterienstämme untereinander verglichen, indem paarweise Vergleiche zwischen den einzelnen Proteinsequenzen durchgeführt und diese dementsprechend zu Clustern von vermutlich orthologen Genen (CLOG) gruppiert wurden. Mithilfe dieser Cluster konnte zwischen Kern-, gemeinsamen und einzigartigen Genen unterschieden werden und es wurde gezeigt, dass die Mehrzahl der Gene nur jeweils in einem einzigen cyanobakteriellen Genom auftreten. Im Gegensatz dazu sind Gene mit einer metabolischen Funktion innerhalb des Kern-Genoms, welches alle analysierten Stämme gemeinsam haben, stark überrepräsentiert. Eine genauere Betrachtung der metabolischen Diversität innerhalb des zentralen Kohlenstoffstoffwechsels zeigt, dass Teile des metabolischen Netzwerkes sehr stark konserviert, aber auch andere Stoffwechselwege stark fragmentiert sind.

Zum Abschluss dieser Arbeit wurde unter Zuhilfenahme der Flussbilanzanalyse die Synthese von neun unterschiedlichen Biokraftstoffen in Cyanobakterien, beziehungsweise am speziellen Beispiel von *Synechocystis* sp. PCC 6803, analysiert. Dabei wurden die Synthesen der Biokraftstoffe charakterisiert und das Zusammenspiel zwischen der Biokraftstoffsynthese und dem Zellwachstum genauer untersucht. Abschließend wird unter Verwendung der Variabilitätsanalyse beschrieben, wie Schlüsselreaktionen beziehungsweise potenzielle Zielgene für eine genetische Manipulation identifiziert werden können.

Vorwort

Diese Arbeit widmet sich der systematischen Beschreibung und Analyse des cyanobakteriellen Stoffwechsels. Der Fokus liegt dabei auf der Rekonstruktion und Analyse genomskaliger metabolischer Netzwerke von phototrophen Mikroorganismen. Die Rekonstruktion genomskaliger metabolischer Netzwerke eröffnet die Möglichkeit einer systematischen Sammlung des Wissens über den Stoffwechsel des jeweiligen Organismus und ist ein guter Ausgangspunkt, um weitere Diskussionen anzuregen, zum Beispiel in Form von experimentellen Messungen, um dieses Wissen zu vermehren und zu verfeinern.

Ein weiterer Aspekt dieser Arbeit ist der Vergleich verschiedener Cyanobakterienstämme untereinander. Dieser ermöglicht ein besseres Verständnis über das Phylum der Cyanobakterien und speziell auch über den phototrophen Stoffwechsel, um zum Beispiel diese Wissensgrundlage später auf höher entwickelte Organismen übertragen zu können. Die Thesen und Behauptungen, die in dieser Arbeit aufgestellt werden, sollen dazu anregen, durch experimentelle Messungen verifiziert zu werden und neue Perspektiven auf den Stoffwechsel zu erschaffen. Der Leser wird zu diesem Zweck durch sechs Kapitel geleitet, die aufeinander aufbauen.

Das erste Kapitel dieser Arbeit gibt eine kurze Einleitung zu den Cyanobakterien, speziell zu den zwei Stämmen *Synechocystis* sp. PCC 6803 (*Synechocystis*) und *Synechococcus elongatus* PCC 7942 (*Synechococcus*), von denen jeweils im Verlauf der Arbeit ein genomskaliges metabolisches Netzwerk rekonstruiert wurde. Überdies werden kurz 14 weitere Cyanobakterienstämme vorgestellt, die für eine Vergleichsanalyse auf Ebene des Genoms und speziell des Stoffwechsel untersucht wurden. Zudem wird ein Überblick über metabolische Netzwerke im Allgemeinen gegeben inklusive deren Rekonstruktion und Modellierung. Abschließend erhält der Leser noch eine kurze Einführung zu Biokraftstoffen und wie insbesondere Cyanobakterien als Biokraftstoffproduzenten eingesetzt werden könnten.

Im zweiten Kapitel wird genauer auf die Rekonstruktion des genomskaligen metabolischen Netzwerkes von *Synechocystis* sp. PCC 6803 und dessen Analyse mithilfe der Flussbilanzanalyse eingegangen. Hierbei steht vor allem die Untersuchung des phototrophen Stoffwechsels, aber auch der Übergang vom Wachstum in Abwesenheit von Licht zum phototrophen Wachstum im Vordergrund.

Das dritte Kapitel behandelt einen Vergleich von 16 verschiedenen Cyanobakterienstämmen auf Ebene des Genoms. Hierfür werden die einzelnen Gene der jeweiligen Organismen in verschiedene Cluster (CLOGs) aufgeteilt, wobei zwischen Kern-Genen, gemeinsamen und einzigartigen Genen unterschieden wird. Anschließend wird besonders der Anteil der metabolischen Cluster im Detail untersucht.

Im darauffolgenden vierten Kapitel wird die Analyse eines genomskaligen metabolischen Modells von *Synechococcus elongatus* PCC 7942 beschrieben, dessen Rekonstruktion durch die vorangegangene Vergleichsanalyse aus dem dritten Kapitel erleichtert wurde. Speziell werden dabei die Unterschiede zum vorher untersuchten metabolischen Netzwerk von *Synechocystis* sp. PCC 6803 herausgearbeitet und dessen Auswirkungen analysiert.

Das fünfte Kapitel bezieht sich auf eine praktische Anwendung der Flussbilanzanalyse metabolischer Netzwerke, in diesem spezifischen Fall der Synthese von Biokraftstoffen mithilfe von Cyanobakterien. Zu diesem Zweck werden die Synthesen von neun unterschiedlichen Biokraftstoffen im metabolischen Netzwerk von *Synechocystis* sp. PCC 6803 analysiert und diskutiert.

Im sechsten und letzten Kapitel dieser Arbeit wird abschließend noch ein zusammenfassendes Fazit über die gesamte Arbeit gezogen und dem Leser ein kurzer Ausblick auf weitere Anwendungen und Projekte gegeben.

Als kurze Information sei angemerkt, dass für die gesamte Thematik dieser Arbeit viele Fachtermini aus dem Englischen stammen. In dieser Arbeit wurde es so gehandhabt, dass wenn eine sinnvolle deutsche Übersetzung des Fachterms existiert, diese genutzt wird, falls aber kein äquivalentes Wort im deutschen Sprachgebrauch existiert, der englische Fachbegriff übernommen wurde. Zudem sind die Beschriftungen einiger Abbildungen ebenfalls in englischer Sprache gegeben, da diese direkt aus der Fachliteratur stammen.

Inhaltsverzeichnis

1. Einleitung	1
1.1. Cyanobakterien	2
1.1.1. <i>Synechocystis</i> sp. PCC 6803	2
1.1.2. <i>Synechococcus elongatus</i> PCC 7942	3
1.1.3. Weitere Cyanobakterienstämme	4
1.2. Systembiologie und metabolische Netzwerke	6
1.2.1. Rekonstruktion metabolischer Netzwerke	7
1.2.2. Modellierung von genomskaligen metabolischen Netzwerken	8
1.3. Biokraftstoffe	10
1.3.1. Cyanobakterien als Biokraftstoffproduzenten	11
1.4. Zielstellung dieser Arbeit	12
2. Das metabolische Netzwerk von <i>Synechocystis</i> sp. PCC 6803	13
2.1. Material und Methoden	15
2.1.1. Computergestützte Modellierung von metabolischen Netzwerken	15
2.1.2. Flussbilanzanalyse	15
2.1.3. Rekonstruktion metabolischer Netzwerke	19
2.1.4. Rekonstruktion des metabolischen Netzwerkes von <i>Synechocystis</i> sp. PCC 6803	21
2.1.5. Annahmen für das metabolischen Netzwerk von <i>Synechocystis</i> sp. PCC 6803	23
2.1.6. Die Synthese von Glycin und Serin	24
2.1.7. Formale Eigenschaften des metabolischen Netzwerkmodells	24
2.1.8. Kriterien für die Modellierung des metabolischen Netzwerkes	25
2.1.9. Die Biomassezielfunktion	27
2.1.10. Zeitaufgelöste Flussbilanzanalyse	28

2.2. Ergebnisse und Diskussion	31
2.2.1. Das metabolische Netzwerk von <i>Synechocystis</i> sp. PCC 6803 . . .	31
2.2.2. Vereinbarkeit bestehender Rekonstruktionen	33
2.2.3. Flussbilanzanalyse des Netzwerkes von <i>Synechocystis</i> sp. PCC 6803	34
2.2.4. Optimale Flussverteilung für das phototrophe Wachstum	36
2.2.5. Die Oxygenase-Reaktion von RuBisCO und die Photorespiration	40
2.2.6. Optimale Flussverteilung bei Abwesenheit von Licht	43
2.2.7. Übergang von der Nutzung von Speicherstoffen hin zum phototrophen Wachstum	45
2.2.8. Der cyanobakterielle Glyoxylat-Shunt	46
2.2.9. Der TCA-Zyklus der Cyanobakterien	49
2.2.10. Der zyklische Fluss durch den TCA-Zyklus in der optimalen Flussverteilung	53
2.2.11. Zeitliche Koordination des phototrophen Stoffwechsels	54
2.3. Schlussfolgerungen	59
3. Die Diversität des cyanobakteriellen Stoffwechsels	63
3.1. Material und Methoden	65
3.1.1. Auswahl der Cyanobakterienstämme	65
3.1.2. Cluster von vermutlich orthologen Genen (CLOG)	66
3.1.3. Anreicherung der GO-Annotation	68
3.1.4. Phylogenetische Analyse	68
3.1.5. Zuweisung von metabolischen Funktionen zu den CLOGs	68
3.2. Ergebnisse und Diskussion	70
3.2.1. Genomanalyse und Cluster von vermutlich orthologen Genen . .	70
3.2.2. Das cyanobakterielle Kern- und Pan-Genom	72
3.2.3. Phylogenetische Übereinstimmungen zwischen den 16 Cyanobakterienstämmen	74
3.2.4. Das metabolische Netzwerk der Cyanobakterien ist stark konserviert	77
3.2.5. Die Diversität des cyanobakteriellen Stoffwechsels	78
3.2.6. Der Stoffwechsel von unterschiedlichen Speicherstoffen bei Cyanobakterien	80
3.2.7. Die Diversität des Zentralstoffwechsels unter den Cyanobakterien	82
3.3. Schlussfolgerungen	87
4. Das metabolische Netzwerk von <i>Synechococcus elongatus</i> PCC 7942	89
4.1. Material und Methoden	91
4.1.1. Rekonstruktion des metabolischen Netzwerkes von <i>Synechococcus elongatus</i> PCC 7942	91
4.1.2. Flussbilanzanalyse des metabolischen Netzwerkes von <i>Synechococcus elongatus</i> PCC 7942	92
4.2. Ergebnisse und Diskussion	93
4.2.1. Das metabolische Netzwerk von <i>Synechococcus elongatus</i> PCC 7942	93

4.2.2.	Der Zentralstoffwechsel von <i>Synechococcus elongatus</i> PCC 7942 im Vergleich zu <i>Synechocystis</i> sp. PCC 6803	93
4.2.3.	Flussbilanzanalyse des metabolischen Netzwerkes von <i>Synechococcus elongatus</i> PCC 7942	96
4.2.4.	Verifizierung des metabolischen Netzwerkes von <i>Synechococcus elongatus</i> PCC 7942	99
4.3.	Schlussfolgerungen	101
5.	Modellierung verschiedener Biokraftstoffsynthesen in Cyanobakterien	105
5.1.	Material und Methoden	107
5.1.1.	Integration von Biokraftstoffsynthesen in das metabolische Netzwerk von <i>Synechocystis</i> sp. PCC 6803	107
5.1.2.	Parameter und Berechnung der stöchiometrischen Eigenschaften für die Biokraftstoffsynthesen	109
5.1.3.	Trade-off und Variabilitätsanalyse der Biokraftstoffsynthesen . .	110
5.1.4.	Synergetische Effekte der Biokraftstoffsynthesen	110
5.2.	Ergebnisse und Diskussion	112
5.2.1.	Stöchiometrische Modellierung von verschiedenen Biokraftstoffsynthesen	112
5.2.2.	Trade-off zwischen Wachstum und Biokraftstoffsynthese	113
5.2.3.	Synergetische Effekte bei der Biokraftstoffsynthese	115
5.2.4.	Variabilitätsanalyse zur Identifikation von Schlüsselreaktionen bei der Biokraftstoffsynthese	117
5.3.	Schlussfolgerungen	121
6.	Resümee und Ausblick	123
A.	Netzwerk von <i>Synechocystis</i> sp. PCC 6803	127
.	Literaturverzeichnis	135
.	Danksagung	151

Abbildungsverzeichnis

1.1. Elektronenmikroskopische Aufnahme von <i>Synechocystis</i> sp. PCC 6803 .	3
1.2. Elektronenmikroskopische Aufnahme von <i>Synechococcus elongatus</i> PCC 7942	3
2.1. Grundlagen der <i>Constraint</i> -basierten Modellierung	16
2.2. Schematischer Ablauf einer Netzwerkrekonstruktion	21
2.3. Überblick über die Attribute des metabolischen Netzwerkes von <i>Synechocystis</i> sp. PCC 6803	32
2.4. Optimale Flussverteilung für <i>Synechocystis</i> sp. PCC 6803 unter phototrophen Bedingungen	37
2.5. Die maximale Wachstumsrate in Abhängigkeit zur Lichtintensität und HCO_3^- -Aufnahme	38
2.6. Die Photorespiration und der Glyoxylat-Stoffwechsel	41
2.7. Optimale Flussverteilung für <i>Synechocystis</i> sp. PCC 6803 bei Abwesenheit von Licht	44
2.8. Übergang von der Nutzung von Speicherstoffen hin zum phototrophen Wachstum	46
2.9. Experimentelle Überprüfung des Glyoxylat-Shunts	48
2.10. Alternative Möglichkeiten des TCA-Zyklus	50
2.11. Auswirkungen des TCA-Zyklus auf das optimale Wachstum	53
2.12. Phasensortierte Expressionsprofile von metabolischen Genen in <i>Synechocystis</i> sp. PCC 6803	55
2.13. Zeitabhängige Biomassezielfunktion	57
2.14. Zeitlicher Verlauf des metabolischen Flusses einiger ausgewählter Enzyme	58
3.1. Anzahl der Gene pro CLOG	70

3.2. Verteilung der CLOGs innerhalb der 16 untersuchten Cyanobakterienstämme	71
3.3. Das cyanobakterielle Kern- und Pan-Genom	72
3.4. Phylogenetische Analyse der 16 untersuchten Cyanobakterienstämme .	75
3.5. Verteilung von metabolischen CLOGs innerhalb der cyanobakteriellen Genome	77
3.6. Vergleich von metabolischen und nicht metabolischen CLOGs	78
3.7. Diversität des cyanobakteriellen Stoffwechsels	79
3.8. Stoffwechseldiagramm des cyanobakteriellen Zentralstoffwechsels . . .	84
4.1. Unterschiede auf Ebene des TCA-Zyklus zwischen <i>Synechocystis</i> sp. PCC 6803 und <i>Synechococcus elongatus</i> PCC 7942	95
4.2. Alternativer Stoffwechselweg zur Glutamat-Dehydrogenase in <i>Synechococcus elongatus</i> PCC 7942	96
4.3. Optimale Flussverteilung für <i>Synechococcus elongatus</i> PCC 7942 unter phototrophen Bedingungen	97
4.4. Optimale Flussverteilung für <i>Synechococcus elongatus</i> PCC 7942 bei Abwesenheit von Licht	98
5.1. Synthesereaktionen der Biokraftstoffe	107
5.2. Trade-off zwischen Wachstum und Biokraftstoffsynthese	114
5.3. Analyse des Synergieeffekts der unterschiedlichen Biokraftstoffe	116
5.4. Variabilitätsanalyse von Reaktionen mit einer abfallenden Flussaktivität für die Synthese von Ethylen	118
5.5. Variabilitätsanalyse des EFE, der PEPC und der SDH für die Synthese von Ethylen	119
A.1. Stoffwechselkarte des Netzwerkes von <i>Synechocystis</i> sp. PCC 6803 . . .	134

Abkürzungen

2PGL	2-Phosphoglycolat
2OGDC	2-Oxoglutarat-Decarboxylase (EC 4.1.1.71)
6PGD	6-Phosphogluconat-Dehydrogenase (EC 1.1.1.44)
6PGL	6-Phosphogluconolactonase (EC 3.1.1.31)
Aca11017	<i>Acaryochloris marina</i> MBIC 11017
ACEP	Acetylphosphat
ACO	Aconitase (EC 4.2.1.3)
AGP	ADP-Glucose-Pyrophosphorylase (EC 2.7.7.27)
CLOG	Cluster von vermutlich orthologen Genen (<i>Cluster of Likely Ortholog Genes</i>)
CoA	Coenzym A
CphA	Cyanophycin-Synthetase (EC 6.3.2.29/30)
CphB	Cyanophycinase (EC 3.4.15.6)
CS	Citrat-Synthase (EC 2.3.3.1)
Cyn51142	<i>Cyanothece</i> ATCC 51142
Cyn8801	<i>Cyanothece</i> PCC 8801
CT	circadiane Zeit (<i>circadian time</i>)
BOF	Biomassezielfunktion (<i>Biomass Objective Function</i>)
BHR	bidirektionale Trefferrate (<i>Bidirectional Hit Rate</i>)
DW	Trockengewicht (<i>Dry Weight</i>)
DCMU	3-(3,4-Dichlorphenyl)-1,1-Dimethylharnstoff
E4P	Erythrose-4-Phosphat
EFE	Ethylen-bildendes Enzym (EC 1.14.11.34)
ENO	Enolase (EC 4.2.1.11)
ETK	Elektronentransportkette
F6P	Fructose-6-Phosphat
FBP	Fructose-1,6-Bisphosphatase (EC 3.1.3.11)

FBA	Flussbilanzanalyse (<i>Flux Balance Analysis</i>)
FBPA	Fructose-1,6-Bisphosphat-Aldolase (EC 4.1.213)
FH	Fumarat-Hydratase (EC 4.2.1.2)
LP	Linearen Programmierung
GABA	γ -Aminobutyrat
GAP	Glyceraldehyd-3-Phosphat
GAPDH	Glyceraldehyd-3-Phosphat-Dehydrogenase (EC 1.2.1.12/59)
GBE	1,4- α -Glucan-verzweigendes Enzym (<i>Glycogen Branching Enzyme</i>) (EC 2.4.1.18)
GDH	Glutamat-Dehydrogenase (EC 1.4.1.4)
Glo7421	<i>Gloeobacter violaceus</i> PCC 7421
GLS	Glutamat-Synthase (EC 1.4.1.14)
GP	Glykogen-Phosphorylase (EC 2.4.1.1)
GPD	Glucose-6-Phosphat-Dehydrogenase (EC 1.1.1.49)
GPI	Glucose-6-Phosphat-Isomerase (EC 5.3.1.9)
GS	Glykogen-Synthase (EC 2.4.1.21)
ICD	Isocitrat-Dehydrogenase (EC 1.1.1.41/2)
ICL	Isocitratlyase (EC 4.1.3.1)
ME	Malatenzym (EC 1.1.1.38)
MDH	Malat-Dehydrogenase (EC 1.1.1.37)
MQO	Malat:Chinon-Oxidoreduktase (EC 1.1.5.4)
MS	Malat-Synthase (EC 2.3.3.9)
Mic843	<i>Microcystis aeruginosa</i> NIES-843
Nos7120	<i>Nostoc</i> sp. PCC 7120
ODE	gewöhnliche Differentialgleichung (<i>Ordinary Differential Equation</i>)
OGDH	2-Oxoglutarat-Dehydrogenase
OOR	2-Oxoglutarat-Oxidoreduktase
OPP	oxidativer Pentosephosphatweg
PEP	Phosphoenolpyruvat
PDH	Pyruvatdehydrogenase
PG3	Glycerat-3-Phosphat
PEPC	PEP-Carboxylase (EC 4.1.1.31)
PEPK	PEP-Carboxykinase (EC 4.1.1.49)
PFK	Phosphofructokinase (EC 2.7.1.11)
PGK	Phosphoglyceratkinase (EC 2.7.2.3)
PGM	Phosphoglucomutase (EC 5.4.2.1)
PhaA	Polyhydroxyalkanoat-spezifische β -Ketothiolase/Acetyl-CoA Acetyltransferase (EC 2.3.1.9)
PhaB	PHA-spezifische Acetoacetyl-CoA-Reduktase (EC 1.1.1.36)
PhaC	Poly-(3-Hydroxyalkanoat)-Synthase-Komponente C (EC 2.3.1.-)
PhaE	Poly-(3-Hydroxyalkanoat)-Synthase-Komponente E (EC 2.3.1.-)
PHB	Poly- β -Hydroxybutyrat
PKL	Phosphoketolase (EC 4.1.2.22)

PPS	PEP-Synthase (EC 2.7.9.2)
PRK	Phosphoribulokinase (EC 2.7.1.19)
ProMED4	<i>Prochlorococcus marinus</i> MED4
Pro9211	<i>Prochlorococcus marinus</i> MIT 9211
Pro9215	<i>Prochlorococcus marinus</i> MIT 9215
PS	Photosystem
PYK	Pyruvatkinase (EC 2.7.1.40)
RPI	Ribose-5-Phosphat-Isomerase (EC 5.3.1.6)
RPE	Ribulose-5-Phosphat-3-Epimerase (EC 5.1.3.1)
RuBisCO	Ribulose-1,5-Bisphosphat-Carboxylase/-Oxygenase (EC 4.1.1.39)
RuBP	Ribulose-1,5-Bisphosphat
SBP	Fructose-1,6-Bisphosphatase/ Sedoheptulose-1,7-Bisphosphatase (EC 3.1.3.37)
SDH	Succinat-Dehydrogenase (EC 1.3.99.1)
SNA	stöchiometrische Netzwerkanalyse
SSADH	Succinat-Semialdehyd-Dehydrogenase (EC 1.2.1.16)
STK	Succinat-Thiokinase (EC 6.2.1.5)
SycJA23	<i>Synechococcus</i> sp. JA-2-3B'a(2-13)
Syc7002	<i>Synechococcus</i> sp. PCC 7002
Syc7803	<i>Synechococcus</i> sp. WH7803
Syc7942	<i>Synechococcus elongatus</i> PCC 7942
Syn6803	<i>Synechocystis</i> sp. PCC 6803
Synechococcus	<i>Synechococcus elongatus</i> PCC 7942
Synechocystis	<i>Synechocystis</i> sp. PCC 6803
TALDO	Transaldolase (EC 2.2.1.2)
TCA-Zyklus	Tricarbonsäurezyklus (<i>Tricarboxylic Acid Cycle</i>)
ThermoBP1	<i>Thermosynechococcus elongatus</i> BP-1
TKT	Transketolase (EC 2.2.1.1)
TPI	Triosephosphat-Isomerase (EC 5.1.3.1)
Trich101	<i>Trichodesmium erythraeum</i> IMS101
X5P	Xylose-5-Phosphat

KAPITEL 1

Einleitung

1.1. Cyanobakterien

Die Cyanobakterien, welche noch bis zum Jahr 1974 auch als Blaualgen bezeichnet wurden, gehören zu den gramnegativen Bakterien. Sie stellen die einzige Gruppe der bisher bekannten Prokaryoten dar, die die Fähigkeit zur oxygenen Photosynthese besitzen und somit in der Lage sind, elementaren Sauerstoff (O_2) zu produzieren [1]. Mithilfe der Photosynthese erzeugen die Cyanobakterien ATP und das Reduktionsäquivalent NADPH, welches sie benutzen, um verschiedene Nährstoffe zu assimilieren und Kohlenstoffdioxid aus der Atmosphäre über den Calvin-Benson-Zyklus zu fixieren. Deswegen spielten Cyanobakterien auch bei der Entwicklung der heutigen Biosphäre vor ca. 3,5 Milliarden Jahren eine entscheidende Rolle [2]. Die damals reduzierte Atmosphäre wurde im Laufe der Zeit mit Sauerstoff angereichert, wodurch eine Basis für eine aerobe Lebensweise auf der Erde geschaffen wurde. Diese ermöglichte die Entstehung einer Vielzahl von Lebewesen und Arten auf der Erde. Auch heute noch tragen die Cyanobakterien einen großen Anteil zur globalen CO_2 -Assimilierung und der Regenerierung von Sauerstoff und Stickstoff bei. Sie werden zudem als die Vorläufer der Plastiden in Pflanzen und Algen (Endosymbiontentheorie) angesehen [3]. Die Morphologie von Cyanobakterien reicht von einzelligen, kokkoiden Organismen bis hin zu mehrzelligen, filamentösen Zellverbindungen. Auf Grund dieser hohen Diversität besitzen Cyanobakterien die Fähigkeit, die unterschiedlichsten Ökosysteme zu besiedeln [4]. Darunter fallen Süß- und Salzwasserhabitate, aber auch heiße Quellen oder sogar vorübergehend feuchte Felsen in der Wüste bis hin zu Polarregionen [5].

In letzter Zeit gewinnt auch die Nutzung von Cyanobakterien für wirtschaftliche Zwecke mehr und mehr an Aufmerksamkeit. Dazu gehören die pharmazeutische Wirkstoffforschung sowie die Produktion von Naturstoffen [6, 7]. Insbesondere besitzen sie durch ihre Fähigkeit der direkten Umwandlung von atmosphärischem CO_2 zu Biomasse und organischen Verbindungen mithilfe von Sonnenlicht ein erhebliches Potenzial als neuartige und erneuerbare Quelle für die Erzeugung von Bioenergie beziehungsweise Biokraftstoffen [8–11]. Außerdem sind Cyanobakterien und Mikroalgen, soweit bekannt, die einzigen Organismen, welche in der Lage sind, sowohl oxygene Photosynthese als auch Wasserstoffproduktion zu betreiben. Die Produktion von H_2 durch phototrophe Mikroorganismen spielt neben der Produktion von Biokraftstoffen (siehe Kapitel 5) eine wichtige Rolle bei der Suche nach alternativen Prozessen zur Erzeugung von erneuerbaren Energien.

1.1.1. *Synechocystis* sp. PCC 6803

Synechocystis sp. PCC 6803 (*Synechocystis*) (Abbildung 1.1) ist ein einzelliges, nicht stickstofffixierendes Cyanobakterium. Unter den verschiedenen Cyanobakterienstämmen ist *Synechocystis* einer der am besten studierten Modellorganismen. Der Stamm wurde zuerst aus einem Süßwassersee in Kalifornien isoliert und wurde 1968 in die *Pasteur Culture Collection* (PCC) übernommen [12]. Sein Genom wurde bereits 1996 als erster photosynthetischer Organismus vollständig sequenziert [13]. Es hat eine

Größe von insgesamt 3,96 Mb und setzt sich aus einem zirkulären Chromosom mit 3,57 Mb und sieben Plasmiden mit insgesamt 390 kb zusammen.

Durch die frühe Sequenzierung existiert mittlerweile eine umfangreiche Sammlung von genomischen, biochemischen und physiologischen Daten für *Synechocystis*, welche es zu einem optimalen Modellorganismus für Analysen photosynthetischer Prozesse macht. Auch seine natürliche Kompetenz [15] extrazelluläre DNA aufzunehmen und in sein Genom zu integrieren, machen das Bakterium leicht empfänglich für genetische Modifikationen. *Synechocystis* wächst phototroph, einige Stämme sind aber auch in der Lage heterotroph, unter der Zugabe von Glukose, zu wachsen [16, 17]. Diese Charakteristika machen *Synechocystis* zu einem idealen Ausgangspunkt, um mithilfe einer systemischen Analyse Einblicke in den phototrophen Stoffwechsel zu erhalten, um dessen Eigenschaften zu beschreiben.

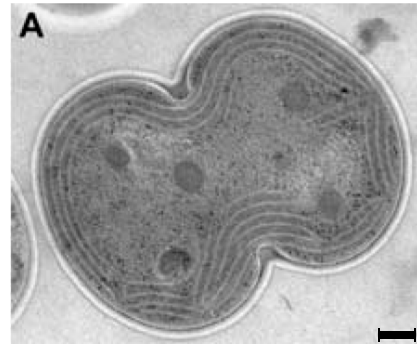


Abbildung 1.1.: Elektronenmikroskopische Aufnahme einer sich teilenden *Synechocystis* sp. PCC 6803-Zelle (Maßstab: 200 nm) (Abbildung aus Kerfeld et al. (2005) [14]).

1.1.2. *Synechococcus elongatus* PCC 7942

Synechococcus elongatus PCC 7942 (*Synechococcus*) (früher als *Anacystis nidulans* bezeichnet; Abbildung 1.2) gehört zu der Gruppe der Süßwasser-Cyanobakterien und besitzt ein obligat photoautotrophes Wachstum [18–20]. Bis heute wurden zwei Stämme der Gattung *Synechococcus elongatus* sequenziert, (PCC 6301 und PCC 7942) [21, 22]. Die Genomgröße von *Synechococcus* umfasst hierbei ca. 2,7 Mb zuzüglich eines Plasmides mit 46 kb. In Bezug auf genetische Manipulationen konnte für *Synechococcus* als erstes Cyanobakterium überhaupt gezeigt werden, dass eine Transformation rein durch externe Zugabe von externer DNA möglich ist [24].

Ferner wird *Synechococcus* als „einfachster“ Modellorganismus für Studien der circadianen Uhr in Prokaryoten angesehen [25]. Es konnte gezeigt werden, dass viele Gene in *Synechococcus* einen circadianen Expressionsrhythmus besitzen [26]. Durch groß angelegte Inaktivierungsexperimente konnten auch die entsprechenden Genloci, welche für das circadiane Verhalten und die circadianen Rhythmen der Genexpression wesentlich sind, identifiziert werden [27]. Es stellte sich heraus, dass die circadianen Rhythmen der Genexpression durch die drei Uhrgene *kaiA*, *kaiB*, und *kaiC* reguliert werden [26]. In der Literatur sind auch schon zahlreiche Modelle zu finden, die die-

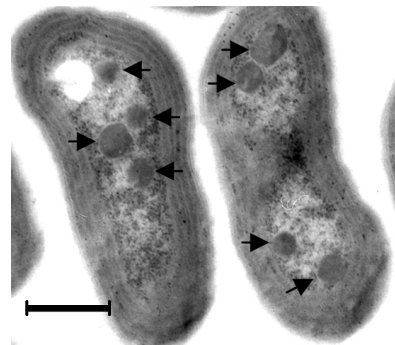


Abbildung 1.2.: Elektronenmikroskopische Aufnahme von *Synechococcus elongatus* PCC 7942. Die Carboxysomen der Zelle sind durch Pfeile hervorgehoben (Maßstab: 500 nm) (Abbildung aus Rae et al. (2012) [23]).

sen Prozess beschreiben [28–30]. *Synechococcus* wächst autotroph relativ schnell, besitzt die Fähigkeit zu natürlicher Kompetenz und lässt sich einfach kultivieren, weshalb es für biotechnologische Anwendungen als ein idealer Kandidat gut geeignet ist. Unter anderem wurden Studien zur Produktion von kurzkettigen Alkoholen, wie Ethanol und Isobutanol, mit *Synechococcus* veröffentlicht [8, 9].

1.1.3. Weitere Cyanobakterienstämme

Neben *Synechocystis* und *Synechococcus* sind noch weitere Stämme von Cyanobakterien aufgrund ihrer hohen Vielfaltigkeit von großem wissenschaftlichen Interesse. Für eine Analyse der Diversität auf der Genomebene (siehe Kapitel 3) wurde eine Auswahl von 16 Stämmen (inklusive *Synechocystis* und *Synechococcus*), die gattungsübergreifend Organismen von allen bekannten Umgebungen integriert, für diese Arbeit getroffen. Dabei wurde darauf Wert gelegt, die Stämme so auszuwählen, dass sie die bekannte genomische und metabolische Vielfalt des cyanobakteriellen Phylums repräsentieren. Darunter fallen einige Süßwasserstämme sowie marine und stickstofffixierende, also auch zwei thermophile Stämme, welche ursprünglich aus heißen Quellen isoliert wurden. Neben *Synechocystis* und *Synechococcus* gehören zu den 16 analysierten Stämmen:

Acaryochloris marina MBIC 11017 (Aca11017)

Aca11017 gehört zu den marinen Cyanobakterien und ist bisher der einzig bekannte Mikroorganismus, der hauptsächlich Chlorophyll *d* als photosynthetisches Pigment verwendet [31].

Cyanothece ATCC 51142 (Cyn51142)

Cyn51142 ist ein einzelliges, stickstofffixierendes, marines Cyanobakterium und war der erste Organismus der Gattung *Cyanothece*, welcher sequenziert wurde [32].

Cyanothece PCC 8801 (Cyn8801)

Cyn8801 ist ein weiteres Cyanobakterium der Gattung *Cyanothece* und somit auch einzellig und stickstofffixierend. Jedoch wurde dieser Stamm aus einer Süßwasserquelle isoliert.

Gloeobacter violaceus PCC 7421 (Glo7421)

Glo7421 ist ein einzelliges, stäbchenförmiges Süßwassercyanobakterium, welches Kolonien bildet, die von einem klebrigen Beutel umhüllt sind. Es wurde als erstes an Kalkfelsen in den Schweizer Alpen isoliert [33].

Microcystis aeruginosa NIES-843 (Mic843)

Mic843 ist ein kolonienbildendes Cyanobakterium, welches im Süßwasser lebt. Es produziert schädliche Toxine, die freigesetzt werden und für Menschen und Tiere gesundheitsschädigend sind [34].

Nostoc sp. PCC 7120 (Nos7120)

Nos7120, auch bekannt unter *Anabaena* sp. PCC 7120, ist ein filamentbildendes Süßwassercyanobakterium und gilt als Modellorganismus für die Bildung von Heterocysten und deren zugehörige Genregulation [35].

Prochlorococcus marinus MED4 (ProMED4), *Prochlorococcus marinus* MIT 9211 (Pro9211) und *Prochlorococcus marinus* MIT 9215 (Pro9215)

ProMED4, Pro9211 und Pro9215 sind drei Stämme der Gattung *Prochlorococcus*, welche zu den einzelligen, marinen Cyanobakterien gehören. *Prochlorococcus* zählt zu den kleinsten bis jetzt bekannten photosynthetisch-aktiven Mikroorganismen und ist in hohen Konzentrationen in weiten Bereichen der Ozeane beheimatet [36].

Synechococcus sp. JA-2-3B'a(2-13) (SycJA23)

SycJA23 ist ein einzelliges, thermophiles Cyanobakterium, welches in der Lage ist, Stickstoff zu fixieren. Es wurde aus cyanobakteriellen Matten in heißen Quellen des Octopus Springs im Yellowstone Nationalpark in den USA isoliert. In diesem Habitat herrschen typischerweise Temperaturen von 50 – 70 Grad Celsius [37].

Synechococcus sp. PCC 7002 (Syc7002)

Syc7002 ist ein einzelliges Cyanobakterium, welches ursprünglich zuerst an der Küste von Puerto Rico isoliert wurde. Es ist sehr lichttolerant und gehört zu den am schnellsten wachsenden Cyanobakterien [38].

Synechococcus sp. WH7803 (Syc7803)

Syc7803 ist ein einzelliges, marines Cyanobakterium, welches zur Gattung *Synechococcus* gehört. *Synechococcus* ist über eine große Spanne von unterschiedlichen Breitengraden weltweit in Küstenbereichen oder Flussmündungen verbreitet [39].

Thermosynechococcus elongatus BP-1 (ThermoBP1)

ThermoBP1 ist ein thermophiles, stäbchenförmiges und einzelliges Süßwassercyanobakterium, welches in heißen Quellen zu finden ist. Es gilt als Modellorganismus für Untersuchungen des Photosyntheseapparates z.B. durch Kristallisation der photosynthetischen Proteine [40].

Trichodesmium erythraeum IMS101 (Trich101)

Trich101 ist ein filamentöses, marines und stickstofffixierendes Cyanobakterium. Es ist in nährstoffarmen, tropischen und subtropischen Ozeanbereichen zu finden. Es wird angenommen, dass diese Cyanobakterien bis zu 40 % der gesamten Stickstofffixierung in den Ozeanen verrichten [41].

1.2. Systembiologie und metabolische Netzwerke

„Der Forschungsansatz der Systembiologie zielt darauf ab, zu einem umfassenden quantitativen Verständnis der dynamischen Interaktionen zwischen den Bausteinen und Komponenten eines biologischen Systems zu gelangen, um das Verhalten des Systems als Ganzes zu verstehen und Vorhersagen zu ermöglichen. Zur Erreichung dieses Ziels werden mathematische Konzepte auf biologische Systeme angewandt. Von zentraler Bedeutung ist hierbei ein iterativer Prozess zwischen Laborexperiment und Modellierung im Computer.“¹

Schon in der ersten Hälfte des vergangenen Jahrhunderts gab es biologische Teilgebiete (zum Beispiel in der Ökologie), in welchen Systemanalysen beziehungsweise integrative Studien durchgeführt worden sind. Der Begriff der „Systembiologie“ im Kontext der modernen Biologie wurde zuerst Ende der 1990er Jahre eingeführt [42, 43]. Der Hauptgrund hierfür war die Einführung der funktionalen Genomik, welche den Zugang zu einer großen Anzahl an hochqualitativen Daten ermöglichte. Hinzu kam noch die technische Weiterentwicklung von Computern, welche die Auswertung dieser Datenmengen überhaupt in angemessener Zeit ermöglichten und vereinfachten. Dies hatte zur Folge, dass nun komplexere und realistischere Modelle zu biologischen Fragestellungen möglich waren. Zu dieser Zeit wurden auch in der Gruppe von Tomita [44] die ersten Ansätze für ein quantitatives Modell des gesamten Metabolismus einer einzelnen Zelle entwickelt.

Allgemein stellt die Rekonstruktion von Modellen einen fundamentalen Schritt bei systembiologischen Analysen dar. Dabei werden alle am Prozess beteiligten biologischen Komponenten, die Wechselwirkungen, Reaktionen und Interaktionen zwischen ihnen sowie, falls vorhanden, Werte für Konzentration oder Lokalisierung mit einbezogen. Die Gesamtheit aller Komponenten inklusive ihrer Beziehungen untereinander bilden ein Netzwerk. Im Falle eines metabolischen Netzwerkes wäre zum Beispiel jeder Metabolit eine Komponente des Netzwerkes und eine Reaktion wandelt einen Metaboliten zu einem oder mehreren anderen Metaboliten beziehungsweise Produkten um.

Für die Rekonstruktion eines Netzwerkes kann jegliche mögliche Information aus der Literatur über Genomsequenzinformationen bis hin zu experimentell erworbenen Daten verwendet werden [45]. Somit hat die Netzwerkrekonstruktion eine zentrale Bedeutung für die späteren Analysen des Systems, denn die Ergebnisse, die ein Modell liefert, hängen stark von der Qualität der Eingangsdaten ab. Fehlerhafte Eingangsdaten mit daraus folgenden Fehlern in der Rekonstruktion, können zu Fehlinterpretationen und falschen Schlussfolgerungen führen [46, 47].

Die Rekonstruktion von qualitativ hochwertigen genomskaligen metabolischen Modellen beziehungsweise Netzwerken [48] stellt somit eine elementare Grundlage für die Forschung im Bereich der Systembiologie dar.

¹(BMBF Bekanntmachung 2007, <http://www.bmbf.de/foerderungen/10604.php>)

1.2.1. Rekonstruktion metabolischer Netzwerke

Die Rekonstruktion eines metabolischen Netzwerkes besteht aus einer Ansammlung von metabolischen Reaktionen und deren zugehörigen Spezies beziehungsweise Substraten, die in einen funktionalen Zusammenhang gesetzt werden. Diese ermöglicht die Verwendung von mathematischen Simulationen, um unterschiedliche Szenarien von Stoffwechselflussverteilungen darzustellen und wenn möglich, zusätzliche Vorhersagen zu treffen.

Für die Reaktionen werden lediglich die Reaktionsrichtung, also die thermodynamischen Eigenschaften und die exakte Stöchiometrie der Substrate und Produkte benötigt. Die größte Herausforderung liegt jedoch darin, die Gesamtheit aller Reaktionen eines Organismus zu sammeln und zu erfassen. Diesbezüglich kann man auf zwei Arten von Informationsquellen zurückgreifen: Einmal auf gemessene Daten aus der Literatur beziehungsweise aus Datenbanken. Im Laufe der letzten Jahrzehnte ist es Wissenschaftlern gelungen, biochemische Reaktionen immer besser zu charakterisieren. Verschiedene Parameter wie die genaue Stöchiometrie der einzelnen Cofaktoren oder die molekulare Aktivität wurden gemessen und definiert und spezielle Bindungsmechanismen und Reaktionsabläufe entschlüsselt [49]. Dies geschah speziell für enzymatische Reaktionen im Stoffwechsel verschiedenster Organismen. Diese Informationen sind in großen Online-Datenbanken gesammelt, wie zum Beispiel BRENDA [50], KEGG [51] oder BioCyc [52].

Eine weitere Möglichkeit der Datenakquise besteht darin, auf Daten zurückzugreifen, die direkt aus der Genomsequenz eines Organismus abgeleitet werden können [53, 54]. Dabei werden unter Zuhilfenahme von unterschiedlichen Werkzeugen, wie dem BLAST- [55] oder dem CLustalW-Algorithmus [56], Enzyme über eine Sequenzähnlichkeit zu anderen Organismen identifiziert und annotiert. Zum Beispiel werden bei neu sequenzierten Organismen unterschiedliche Suchalgorithmen verwendet, um Sequenzen zu finden, die eine hohe Ähnlichkeit zu bereits bekannten Enzymen besitzen. Hierbei wird die Vereinfachung oder Verallgemeinerung getroffen, dass eine hohe Ähnlichkeit auf Sequenzebene eine gleiche oder ähnliche enzymatische Funktion impliziert.

Obwohl bei diesem Schritt ein automatisierter Ablauf durch unterschiedliche Methoden leicht möglich wäre, sind die daraus resultierenden Netzwerkrekonstruktionen schnell fehlerbehaftet und lückenhaft. Deshalb ist eine anschließende manuelle Kuration, also eine Überprüfung und wenn möglich Vervollständigung der einzelnen Stoffwechselwege, ein essentieller Schritt der Rekonstruktion von metabolischen Netzwerken. Dies ist eine sehr schwierige und zeitaufwändige Aufgabe, die eine Zeitspanne von mehreren Monaten bis zu einigen Jahren umfassen kann [48].

Es sollte hier angemerkt werden, dass es oft sinnvoll ist, zusätzliche Hilfsreaktionen zum metabolischen Netzwerk hinzuzufügen, die auch Hilfsspezies, also keine konkreten biochemischen Verbindungen, verwenden. Ein Beispiel dafür ist die Biomassereaktion [57]. Falls das stöchiometrische Verhältnis von jedem Bestandteil der Biomasse bekannt ist, kann die Biomassereaktion als eine Reaktion definiert werden, welche eine externe Hilfsspezies, in den meisten Fällen die Biomasse, produziert und dabei je-

de notwendige Einzelkomponente der Biomasse im entsprechenden Verhältnis konsumiert. Die Biomassereaktion wird bei metabolischen Netzwerken oft als Zielfunktion definiert, um zum Beispiel bei einer Flussbilanzanalyse (*Flux Balance Analysis*) (FBA) das Zellwachstum zu modellieren [58].

1.2.2. Modellierung von genomskaligen metabolischen Netzwerken

Regulatorische und metabolische Netzwerke sind die beiden wichtigsten Systeme in der heutigen Systembiologie [59]. Das Zusammenspiel dieser beiden Systeme ist von großer Bedeutung [60]. Dennoch wird bei der Modellierung von metabolischen Netzwerken die Regulation von Reaktionen meist außen vorgelassen. Der Grund hierfür liegt darin, dass genomskalige metabolische Netzwerke sehr viel leichter rekonstruiert und analysiert werden können als genomskalige Genregulationsnetzwerke. Dennoch haben sich auch genomskalige metabolische Netzwerke ohne Vorhandensein der Regulationsebene in der Praxis bewährt. Es konnte gezeigt werden, dass die Vorhersagen für intrazelluläre Flüsse und Wachstumsraten solcher Netzwerke auch für verschiedene Organismen mit gemessenen Raten und Werten gut übereinstimmen [61, 62].

Gewöhnlich werden bei metabolischen Netzwerken das Verhalten und die Dynamik eines Systems mithilfe von gewöhnlichen Differentialgleichungen (ODE) beschrieben. Die hierbei erstellten ODE-Modelle sind deterministische Modelle, das heißt, wenn der Zustand des Systems zu einem bestimmten Zeitpunkt bekannt ist, ist es auch möglich, das Verhalten des Systems zu jedem anderen Zeitpunkt zu bestimmen [63]. Für die Analyse von genomskaligen Netzwerken sind ODEs aus zweierlei Hinsicht nicht geeignet [49]. Erstens sind für genomskalige metabolische Netzwerke die Mechanismen und Parameter für eine beträchtliche Anzahl von Reaktionen unbekannt. Zweitens werden die Reaktionsmechanismen und Parameter in der Regel durch Analysen bestimmt, wobei einzelne Enzyme isoliert werden, also *in vitro*. Das heißt, auch wenn solche Daten für eine jede Reaktion zur Verfügung stehen, gibt es keine Garantie, dass die Enzyme beziehungsweise Reaktionen auch tatsächlich so in der Zelle, also *in vivo*, ablaufen.

Die stöchiometrische Netzwerkanalyse und die *Constraint*-basierte Modellierung

Ein anderer Ansatz für die Modellierung solcher großen Netzwerke ist die *Constraint*-basierte Modellierung beziehungsweise die stöchiometrische Netzwerkanalyse. Dabei werden statt detaillierter Kinetiken und Parameter aller einzelnen biochemischen Reaktionen ausschließlich die physikalischen und chemischen Randbedingungen für das Verhalten eines Systems festgelegt. So kann zum Beispiel herausgefunden werden, welche Szenarien in einem Modell möglich sind, aber zeitlich aufgelöste beziehungsweise dynamische Vorhersagen des Systemverhaltens in kurzen Zeitabständen sind nicht möglich [64]. Kurz gesagt, gibt es einen Trade-off zwischen dem Informationsgehalt der Ergebnisse und der Menge an Informationen, die man in ein Modell integriert.

Die stöchiometrische Netzwerkanalyse (SNA) wurde 1988 von Clarke entwickelt [65] und bezieht sich auf die Stöchiometrie eines Systems, da diese im Gegensatz zu den

kinetischen Eigenschaften meistens bekannt ist. Sie bietet sich besonders für die mathematische Analyse metabolischer Netzwerke an. Die SNA ist ein mathematischer Ansatz, welcher numerisch leicht lösbar ist und eine hohe Voraussagekraft für die Untersuchung der Eigenschaften von größeren metabolischen Netzwerken besitzt. Zusammen mit einer auf *Constraints*-basierten Modellierung liegt der Vorteil der SNA darin, dass sie nur auf wenigen, aber konkreten Gesetzmäßigkeiten und Annahmen basiert, von denen die meisten für prokaryotische metabolische Netzwerke unter vielen physiologischen Bedingungen gelten.

Bei der *Constraint*-basierten Modellierung sind vorweg Informationen oder Annahmen über die einzelnen Mechanismen und Parameter einer Reaktion nicht erforderlich. Wenn jedoch solche Informationen verfügbar sind, wie zum Beispiel dass ein bestimmter intrazellulärer Fluss experimentell gemessen wurde, können diese Informationen als zusätzliche Einschränkung in das Modell integriert werden [49, 66].

Wie im Vorfeld schon erwähnt, besteht ein metabolisches Netzwerk aus zwei Elementen: den Spezies beziehungsweise Metaboliten und den Reaktionen. Hierbei bestimmt eine Reaktion, welche Metaboliten bei welchem Verhältnis miteinander zu einem oder mehreren anderen Metaboliten reagieren. Eine Systembegrenzung, welche zum Beispiel die Zellmembran darstellt, trennt die internen Metaboliten, welche im Innern des Systems sind und im Rechenmodell berücksichtigt werden, von den extrazellulären Metaboliten, die sich oft außerhalb des Rechenmodells befinden und als externe Metaboliten definiert werden. Reaktionen, die in das System führen oder es verlassen, werden als Austauschflüsse deklariert. Metaboliten, die im Netzwerk verbraucht werden, werden keine Austauschflüsse zugewiesen. Den internen Reaktionen sowie auch den Austauschflüssen können feste Raten oder Grenzwerte zugewiesen werden.

1.3. Biokraftstoffe

Mit dem anhaltenden Wachstum der Weltbevölkerung und der fortschreitenden Industrialisierung steigt parallel der weltweite Verbrauch von Kraftstoffen. Die Entwicklung von Technologien zur langfristigen und nachhaltigen Produktion von erneuerbaren Kraftstoffen auf globaler und lokaler Ebene ist ein bedeutendes Anliegen der heutigen Zeit und auch in Zukunft [67]. Da es unvermeidlich ist, dass fossile Brennstoffe in naher oder ferner Zukunft aufgebraucht sind, ist es eine Herausforderung für die Wissenschaft, alternative Produktionsmöglichkeiten aus erneuerbaren Energien zu entwickeln. Für die Erzeugung von Strom existieren bereits einige CO₂-freie Produktionssysteme, wie zum Beispiel Photovoltaik, Windenergie, Erdwärme und Wellen- und Wasserkraftwerke. Jedoch für die Produktion von Kraftstoffen, welche aktuell 70 % des gesamten Energieverbrauchs darstellen [68], ist es noch nicht möglich, in diesem Ausmaß Kraftstoffe in einer nachhaltigen Art und Weise zu erzeugen.

Fossile Brennstoffe dienen nicht nur als Kraftstoff, sondern auch zur Herstellung von Chemikalien, die für das moderne tägliche Leben benötigt werden. Dies führt zu zahlreichen Problemen, unter anderem zu einem alarmierenden Grad an Verschmutzung unseres Ökosystems [69]. Um uns aus dieser Abhängigkeit von fossilen Brennstoffen zu lösen, ist die Entwicklung nachhaltiger industrieller Verfahren zur Synthese von Kraftstoffen und Chemikalien aus Bioressourcen erforderlich [70–72]. Mehrere neue Technologien wurden entwickelt, um das anhaltende Problem der Verschmutzung der Erdatmosphäre und des daraus resultierenden Treibhauseffektes einzudämmen. Hierbei geht es vor allem darum, die Nutzung von fossilen Brennstoffen einzuschränken und durch die Produktion flüssiger Biokraftstoffe der ersten Generation wie Fettsäureester (Biodiesel) [73], Alkane [74] und höhere Alkohole aus erneuerbaren Quellen [75–77] zu kompensieren. Die meisten dieser Prozesse der Biokraftstoffproduktion haben einen großen Nachteil: sie basieren auf der Nutzung von Mikroorganismen, die unterschiedliche Kohlenstoffquellen, hergestellt aus Landpflanzen, verstoffwechseln. Für die Kohlenstoffquellen dienen vor allem höhere Pflanzen wie Mais, Zuckerrohr, Soja oder Ölpalmen als Grundlage [78, 79].

Damit steht die Biokraftstoffproduktion jedoch im Wettbewerb mit der landwirtschaftlichen Nutzung von Flächen für Nahrungspflanzen und führt somit zu erhöhten Lebensmittelkosten [80, 81]. Zudem führt die Herstellung von Biokraftstoffen aus Landpflanzen (wie zum Beispiel Zuckerrohr) zu großen ökologischen Kosten im Vergleich zur Nutzung von fossilen Brennstoffen. Hier sind vor allem die Abholzung von Regenwäldern, welche durch die Brandrodung eine große Menge an Emissionen von Treibhausgasen und beträchtliche Schäden am Ökosystem verursachen, zu nennen [81]. In neuester Zeit gilt in der Wissenschaft deswegen das Interesse besonders der Nutzung von aquatischen photobiologischen Mikroorganismen, wie Cyanobakterien oder auch Grünalgen, die als biotechnologische Nutzorganismen für die Umwandlung von Sonnenenergie, H₂O und CO₂ in Kohlenwasserstoffkraftstoffe [82] oder deren Vorstufen [83] genutzt werden können.

1.3.1. Cyanobakterien als Biokraftstoffproduzenten

Die direkte Umwandlung von Sonnenenergie in einen flüssigen Kraftstoff durch photosynthetische Mikroorganismen ist eine attraktive Alternative zu fossilen Brennstoffen. Besonders hervorzuheben ist diesbezüglich die Verwendung von Cyanobakterien als Produzenten, da sie mehrere Vorteile besitzen: Das Genom vieler Stämme steht bereits vollständig sequenziert zur Verfügung und es existieren Werkzeuge, die eine genetische Manipulation leicht machen beziehungsweise sie von sich aus natürlich transformierbar sind [24, 84–87]. Sie sind in der Lage, oxygene Photosynthese unter Verwendung von Wasser als Elektronendonator zu betreiben und können bis zu 10 % der Sonnenenergie in Biomasse konvertieren. Damit sind sie in ihrer Leistung Pflanzen und Algen überlegen. Im Vergleich erreichen herkömmliche Pflanzen wie Mais oder Zuckerrohr nur eine Rate von ca. 1 % und Algen bis zu 5 % [88]. Insbesondere besitzen Cyanobakterien eine höhere Wachstumsrate im Vergleich zu Pflanzen (Biomasseproduktion) und können durch ihre natürliche Vielfalt an einer Vielzahl von Standorten zu einer hohen Dichte wachsen, die per se nicht für die Landwirtschaft genutzt werden können. Somit wird der Konflikt zwischen der Nutzung von Ackerflächen für Nahrungsmittel oder zur Herstellung von Biokraftstoffen gelöst. Zudem ist es möglich, die Biokraftstoffe in der Nähe des Kraftstoffverbrauchs zu produzieren und so die Transportkosten zu reduzieren. Des weiteren haben aquatische Systeme gegenüber Landpflanzen den Vorteil, dass diese nicht auf das Vorhandensein von Süßwasser beschränkt sind [89], sondern auch eine Vielzahl anderer Wasserquellen, wie Salz-, Brack-, Meer- und Abwasser [90] nutzen können und nur einfache Nährstoffanforderungen besitzen (vor allem Wasser, Sonnenlicht und CO_2) [87].

In der Literatur wurden bereits mehrere Beispiele von aquatischen Systemen zur direkten Umwandlung von Kohlenstoffdioxid, Wasser und Sonnenlicht in potenzielle Energieträger beschrieben. Darunter fallen zum Beispiel Ethylen [91], Isopren [10], Ethanol [9] und Butanol [82]. Für die Produktion solcher Biokraftstoffe ist es wichtig sicherzustellen, dass eine gute Verträglichkeit zwischen dem Stamm und dem biotechnologischen Endprodukt herrscht. Der gesamte Umwandlungsprozess innerhalb des Produktionsstamms wird stark durch seine metabolischen Kapazitäten limitiert. Insbesondere gilt dies für die Verfügbarkeit von Vorläufermetaboliten und Cofaktoren, aber auch die Toleranzgrenze bezüglich der Toxizität für chemische Zwischenstufen oder des spezifischen Endproduktes können die Produktion einschränken [92]. Andere Faktoren, die auch berücksichtigt werden müssen, sind zum Beispiel das Verfahren zur physischen Trennung des gewünschten chemischen Endprodukts von dem Produktionssystem und die Eignung, Verteilung und Nutzung des Kraftstoffes in den bestehenden Infrastrukturen.

1.4. Zielstellung dieser Arbeit

Das Ziel dieser Arbeit ist eine Beschreibung und Analyse des cyanobakteriellen Stoffwechsels mit dem Fokus speziell auf den Stoffwechsel von *Synechocystis* sp. PCC 6803. Zu diesem Zweck soll zunächst ein metabolisches Netzwerk von *Synechocystis* sp. PCC 6803 rekonstruiert werden, welches mit der Zeit erweitert und verfeinert wird, um eine möglichst getreue Abbildung des Stoffwechsels zu gewähren. Dieses Modell soll zum einen als umfassende Wissenssammlung über den Stoffwechsel und die metabolischen Enzyme dienen und zum anderen unter Verwendung der Flussbilanzanalyse eingehend untersucht werden.

Ein wesentlicher Fokus der Betrachtungen soll dabei auf der Analyse des phototrophen Wachstums bei Anwesenheit von Licht gegenüber dem Stoffwechsel bei Abwesenheit von Licht, der auf der Nutzung von Speicherstoffen beruht, liegen. Die Ergebnisse und Erkenntnisse, die aus dieser systematischen Analyse gewonnen werden, können dazu verwendet werden, um unklare oder noch nicht klar definierte Stellen im metabolischen Netzwerk hervorzuheben und diese zum Beispiel durch neue experimentelle Ansätze zu entschlüsseln. Die neuen Ergebnisse aus solchen Experimenten können wiederum dazu beitragen, das existierende Modell zu validieren und zu verbessern, dabei helfen neue Hypothesen zu formulieren oder durch Abgleich mit dem Netzwerk zu weiteren Diskussionen anzuregen. Zudem soll mithilfe der Flussbilanzanalyse das genomskalige metabolische Netzwerk von *Synechocystis* sp. PCC 6803 für eine Anwendung aus dem Bereich des *Metabolic Engineering* genutzt werden. Zu diesem Zweck soll das metabolische Modell um zusätzliche Synthesereaktionen für verschiedene Biokraftstoffe aus der Literatur erweitert und deren Biosynthesen innerhalb des Netzwerkes eingehend analysiert werden. Neben dem Vergleich der metabolischen Anforderungen und der maximalen theoretischen Energiegewinne der jeweiligen Kraftstoffe untereinander, sollen weiterhin Methoden angewendet werden, die es erlauben mögliche Ziele für eine genetische Manipulation zu identifizieren, die eine effektivere Produktsynthese *in vivo* ermöglichen könnten.

Ein weiteres Ziel dieser Arbeit soll der Vergleich mit dem Stoffwechsel anderer Cyanobakterienstämmen sein. Hierbei soll speziell die Diversität innerhalb des Zentralstoffwechsels und der Speicherstoffe, die für den Stoffwechsel bei Abwesenheit von Licht benötigt werden, im Vordergrund stehen. Da Cyanobakterien die einzigen bekannten Prokaryoten darstellen, die eine oxygene Photosynthese betreiben, können die Ergebnisse und Erkenntnisse aus dem Vergleich mehrerer Cyanobakterienstämmen dazu beitragen, ein besseres Verständnis für den phototrophen Stoffwechsel an sich zu erlangen, aber auch über dessen Entwicklung innerhalb dieses Phylums. Darüber hinaus soll für diese Arbeit die Rekonstruktion eines zweiten metabolischen Netzwerkes eines weiteren Cyanobakterienstammes erfolgen. Das Ziel ist es, dieses wiederum mithilfe der Flussbilanzanalyse zu charakterisieren und zu untersuchen. Dabei soll vor allem der Schwerpunkt auf dem Vergleich zum vorher rekonstruierten metabolischen Netzwerk von *Synechocystis* sp. PCC 6803 liegen, insbesondere auf Basis der Netzwerktopologie und der Analyse des metabolischen Potenzials.

Das metabolische Netzwerk von *Synechocystis* sp. PCC 6803

Dieses Kapitel basiert auf den Veröffentlichungen:

[93] H. Knoop, Y. Zilliges, W. Lockau und R. Steuer (2010). The Metabolic Network of *Synechocystis* sp. PCC 6803: Systemic Properties of Autotrophic Growth. *Plant Physiol.*, 154(1):410–422

[94] H. Knoop, M. Gründel, Y. Zilliges, R. Lehmann, S. Hoffmann et al. (2013). Flux Balance Analysis of Cyanobacterial Metabolism: The Metabolic Network of *Synechocystis* sp. PCC 6803. *PLoS Comput Biol*, 9(6):e1003081

Vorwort

Das folgende Kapitel behandelt die Rekonstruktion und Analyse des genomskaligen metabolischen Netzwerkes des Cyanobakteriums *Synechocystis* sp. PCC 6803. Es werden Details zum Rekonstruktionsprozess des metabolischen Netzwerkes gegeben und die generellen Eigenheiten und Charakteristika des Stoffwechsels von *Synechocystis* sp. PCC 6803 beschrieben. Unter Zuhilfenahme der Flussbilanzanalyse wird der cyanobakterielle Stoffwechsel unter phototrophen Bedingungen und unter Abwesenheit von Licht genauer untersucht. Zu diesem Zweck wird auch die Modellierung eines kompletten Tagesablaufes vorgestellt und das Verhalten einzelner Stoffwechselreaktionen betrachtet, um nähere Schlüsse über die zeitliche Koordination des Stoffwechsels zu ziehen.

Insbesondere behandelt dieses Kapitel die Analyse der Einflüsse von verschiedenen Netzwerktopologien auf die optimale Flussverteilung innerhalb des metabolischen Netzwerkes. Speziell werden die Auswirkungen eines Glyoxylat-Shunts im metabolischen Netzwerk von *Synechocystis* sp. PCC 6803 beschrieben und zusammen mit vorgestellten experimentellen Ansätzen und Daten, gemessen von Marianne Gründel und Yvonne Zilliges, eingehend untersucht und diskutiert.

2.1. Material und Methoden

2.1.1. Computergestützte Modellierung von metabolischen Netzwerken

Im Allgemeinen können Konzentrationsveränderungen der verschiedenen Spezies in einem strukturierten metabolischen Netzwerk über die Zeit als System von Differenzialfunktionen beschrieben werden. Für ein metabolisches Netzwerk N , mit m internen Spezies und n Reaktionen, ist die stöchiometrische Matrix S des Systems eine $(m \times n)$ -Matrix, wobei ein Element S_{ij} der stöchiometrische Koeffizient der Spezies i in der Reaktion j ist. Die Spalten von S werden mit S_1, \dots, S_n bezeichnet. Der Fluss durch die i -te Reaktion wird als v_i definiert. Wenn der Vektor \dot{c} als Änderung der Metabolitkonzentrationen jeder möglichen Flussverteilung deklariert wird, erhält man für jede beliebige Flussverteilung v die Gleichung [95]:

$$S \cdot v = \dot{c}$$

2.1.2. Flussbilanzanalyse

Im Metabolismus eines Organismus laufen Reaktionen normalerweise schnell ab und ebenso verhält es sich auch mit den an den Reaktionen beteiligten Spezies, die schnell umgesetzt werden, im Vergleich zu beispielsweise regulatorischen Schritten. Bei einer Flussbilanzanalyse (*Flux Balance Analysis*) eines metabolischen Netzwerkes wird davon ausgegangen, dass über längere Sicht gesehen, die Metabolitkonzentrationen der internen Spezies und die Reaktionsraten konstant bleiben, dass heißt es existieren keine Netto-Veränderungen der Konzentrationen. Man geht von einer Massenbilanz (*Mass Balance*) aus beziehungsweise genauer von einem Fließgleichgewicht (*Steady State*) [96]. Dies bedeutet für jede stationäre Flussverteilung v^0 gilt $\dot{c} = 0$ (*Steady State*-Annahme). Somit erhält man:

$$S \cdot v^0 = 0$$

Externe Spezies, wie zum Beispiel extrazelluläre Nährstoffe wie Glucose oder Stoffe, die innerhalb der Zelle akkumulieren dürfen, werden bei der Massenbilanz ausgeschlossen. Das homogene System von linearen Gleichungen fordert, dass die Produktion, also die Summe aller positiven Flüsse und der Verbrauch, also die Summe aller negativen Flüsse, einer Spezies über einen gewissen Zeitraum sich aufheben und somit gleich sein müssen. Auf der Ebene des Netzwerkes schränkt die Massenbilanz den Raum der zulässigen Flüsse ein, so dass nur bestimmte Kombinationen von Flusswerten v_i^0 möglich sind. Die triviale Lösung mit $v^0 = 0$ würde ein thermodynamisches Gleichgewicht mit null Freiheitsgraden widerspiegeln und wäre bei der Analyse des metabolischen Netzwerkes einer Zelle nicht von Interesse.

Aus der Linearen Algebra ist bekannt, dass alle möglichen Lösungen dieser Gleichung im sogenannten Nullraum K von S liegen [97]. Die Dimension des Nullraums $\dim(K)$, welche der Anzahl aller linear unabhängigen Lösungen entspricht, ist gleich $n - \text{Rang}(S)$. Der Nullraum K bildet einen konvexen, vielflächigen Kegel (siehe Abbildung 2.1). Die Ränder des Kegels sind durch die sogenannten *Extreme Pathways* [98]

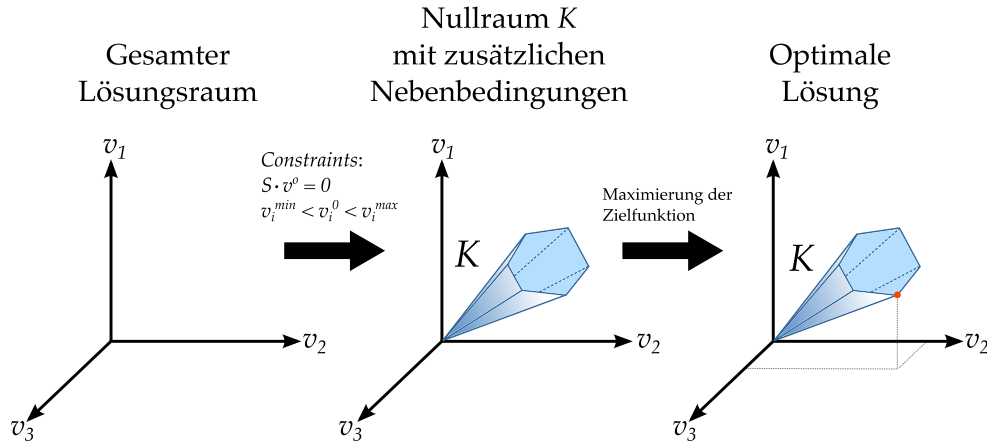


Abbildung 2.1.: Grundlagen der *Constraint*-basierten Modellierung. Ohne Nebenbedingungen kann die Flussverteilung eines biologischen Netzwerkes an jedem Punkt im Lösungsraum liegen. Infolge der Einhaltung der Massenbilanz von S und zusätzlichen Flusseinschränkungen (wie z.B. Reversibilitätseinschränkungen) auf dem Netzwerk, wird ein kleiner Lösungsraum definiert (K). Schlussendlich wird durch die Optimierung einer Zielfunktion mithilfe der FBA eine bestimmte optimale Flussverteilung identifiziert.

definiert und alle Punkte im Inneren des Kegels, also jede mögliche Flussverteilung v^0 eines Genotyps, können als positive Linearkombinationen dieser Wege beschrieben werden.

Für die FBA wird der Lösungsraum der möglichen Flussverteilungen verkleinert, indem zusätzliche Nebenbedingungen auf die einzelnen Flussraten innerhalb des Netzwerkes angewendet werden. Für eine bestimmte Reaktion kann eine Begrenzung der Flussrate beschrieben werden durch:

$$v_i^{min} < v_i^0 < v_i^{max},$$

wobei v_i^{min} und v_i^{max} die untere und obere Grenze des Flusses darstellen. Ferner können durch Setzen einer Grenze auf null, irreversible Reaktionen beschrieben werden. Auch experimentell gemessene Flussraten können auf diese Weise in das Modell einfließen und den Lösungsraum reduzieren. Für eine gemessene Flussrate v_i^m wird die entsprechende Reaktion im Netzwerk mit einer Ungenauigkeit von ε angegeben, so dass gilt:

$$v_i^m - \varepsilon < v_i^0 < v_i^m + \varepsilon$$

Bei $\varepsilon = 0$ wird ein Fluss mit einer bestimmten Rate festgelegt. Hierbei ist zu erwähnen, dass die untere beziehungsweise obere Grenze einer Flussrate für eine bestimmte Reaktion nicht unbedingt ausschlaggebend für die minimal oder maximal möglichen Flusswerte dieser Reaktion sind. Auf Grund der Massenbilanz können Nebenbedingungen für einen bestimmten Fluss auch die Rate von anderen Reaktionen im Netzwerk beeinflussen.

Stöchiometrische und Reversibilitätseinschränkungen bezeichnet Palsson als nicht veränderbare (*non adjustable*) Nebenbedingungen [64], in der Hinsicht, dass sie die intrinsischen Eigenschaften des Systems darstellen. Zu diesen Nebenbedingungen gehören auch die von Beart et al. eingeführten thermodynamischen Nebenbedingungen [99]. Andere Nebenbedingungen, die in der Regel variabel sein können, werden von Palsson als einstellbar (*adjustable*) bezeichnet [64]. Zum Beispiel können diese Bedingungen so gesetzt werden, dass sie bestimmten Umgebungsbedingungen oder biochemischen Bedingungen entsprechen. Hierzu zählen unter anderem regulatorische Rahmenbedingungen [60] oder Metabolitkonzentrationen, die als Nebenbedingungen fungieren [100]. Zusammen mit den nicht verstellbaren Nebenbedingungen können diese die Vorhersage der stöchiometrischen Netzwerkanalyse präzisieren und zuverlässiger machen.

Normalerweise reicht die Einschränkung der Massenbilanz selber nicht aus, um zu einer bestimmten Flussverteilung zu führen. Für die FBA wird deshalb noch angenommen, dass die Zellen über die Zeit ihren Stoffwechsel so entwickelt haben, dass die internen Flüsse der Zelle eine oder mehrere Optimalitätskriterien erfüllen. Das bedeutet, dass für die FBA eine Zielfunktion definiert wird, welche eine Anzahl an möglichen Lösungen im Flussraum identifiziert. Für einzellige Organismen wie Cyanobakterien ist das häufigste Optimalitätskriterium die Maximierung der Biomasse. Da in den meisten Fällen das Netzwerk so rekonstruiert wurde, dass Wachstum und Reproduktion simuliert werden, wird demnach hierfür die Biomassezielfunktion (*Biomass Objective Function*) (BOF) gewählt [58]. Es wird angenommen, dass die metabolischen internen Flüsse so verteilt sind, dass sie eine maximale Ausbeute der Synthese aller metabolischen Komponenten, die für das Zellwachstum benötigt werden, im richtigen Verhältnis ermöglichen.

$$\text{BOF}[v^0] \longrightarrow \text{maximal}$$

Somit ist die resultierende Lösung eine Flussverteilung v^0 , welche die BOF maximiert und die *Steady State*-Bedingung und alle vorher festgelegten Nebenbedingungen berücksichtigt. Eine alternative Flusslösung, welche energetisch ungünstiger sein würde, aber eine schnellere Umwandlung der einzelnen Metaboliten ermöglichen könnte, würde von der FBA nicht gefunden werden [101]. Es konnte schon mehrfach gezeigt werden, dass experimentelle Wachstumsraten von Bakterien für eine Vielzahl von unterschiedlichen Substraten erfolgreich durch eine FBA vorhergesagt werden können [61, 102].

Die Lösung der FBA, also die gesuchte optimale Flussverteilung, wird durch das Lösen der linearen Gleichungen mittels der Linearen Programmierung (LP) gefunden [103]. Ein lineares Programm wird durch eine lineare Zielfunktion und eine Menge von linearen Nebenbedingungen definiert. Dabei kann die Zielfunktion ein Fluss durch eine bestimmte Reaktion oder eine Linearkombination der Flüsse durch die Reaktionen sein. Sind die Nebenbedingungen gut definiert und mit Bedacht gesetzt, ist das Lineare Programm lösbar und eine optimale Lösung wird unter Verwendung einer LP-Solver-Software geliefert.

Anwendungen der Flussbilanzanalyse

Wie beschrieben, kann die FBA die effektivsten und effizientesten Flussverteilungen eines metabolischen Netzwerkes identifizieren, die eine bestimmte Zielfunktion maximieren. Um diese zu berechnen sind unterschiedliche Algorithmen und Software-Tools frei verfügbar.

Im Laufe dieser Arbeit wurden für die meisten Optimierungsanalysen die COBRA-Toolbox (*CO*nstraint-*B*ased *R*econstruction and *A*nalysis) in der Version 2.0.5 (<http://opencobra.sourceforge.net/openCOBRA/Welcome.html>) [104, 105] sowohl als auch der CellNetAnalyzer in der Version 2012.1 (<http://www.mpi-magdeburg.mpg.de/projects/cna/cna.html>) [106] und die FASIMU-Software [107] verwendet.

Die COBRA-Toolbox ist für die quantitative Vorhersage von zellulären Vorgängen konzipiert worden und bedarf zudem, wie der CellNetAnalyzer, einer MATLAB-Umgebung. Außerdem benötigen beide Toolboxen für die Lineare Optimierungsanalyse, um die LP zu lösen, eine separate Installation einer unterstützten LP-Solver-Software. Für diese Arbeit wurden diese Toolboxen in der Umgebung von MATLAB 7.12.634 genutzt, einschließlich der MATLAB-Optimierungs-Toolbox und dem Gurobi Optimizer [108] als lineare LP-Solver-Software. Unter Verwendung der COBRA-Toolbox ist es möglich, eine FBA durchzuführen. Die FBA sucht im Lösungsraum alle Flussverteilungen einer spezifischen Lösung, welche eine vorher definierte Zielfunktion maximieren. Die Schwierigkeit dabei besteht darin, eine geeignete Zielfunktion, wie Maximierung von Biomasse oder Maximierung einer Produktbildungsrate, zu finden [98]. Die stöchiometrischen, metabolischen Netzwerke, die mithilfe der COBRA-Toolbox analysiert werden sollen, müssen vorher manuell oder mit einem anderen systembiologischen Netzwerkmodellierungsprogramm, wie zum Beispiel dem CellNetAnalyzer erstellt und leicht modifiziert werden. Die entsprechenden SBML-Dateien der Netzwerke von *Synechocystis* sp. PCC 6803 und *Synechococcus elongatus* PCC 7942 sind auf der Webseite der elektronischen Version der Veröffentlichung Knoop et al. (2013) [94] beziehungsweise im elektronischen Anhang dieser Arbeit verfügbar.

Unter Verwendung der Flussbilanzanalyse können mittels der COBRA-Toolbox verschiedene selbst- oder vordefinierte Funktionen angewendet werden, um ein gegebenes metabolisches Netzwerk zu untersuchen. Dazu gehören unter anderem:

- Simulieren eines maximalen Wachstums (*Optimal Growth*):
Um das mögliche Maximum der Wachstumsrate zu berechnen, wird als Zielfunktion die Biomassessynthese definiert und über Lineare Optimierung die optimale Flussverteilung berechnet. Hierbei können auch die Flussverteilungen für unterschiedliche Wachstumsbedingungen ermittelt werden.
- Robustheitsanalyse:
Die Robustheitsanalyse erlaubt die Berechnung, wie ein vorher definiertes Ziel (definiert durch die Zielfunktion, zum Beispiel die Wachstumsrate) sich verändert, wenn der Fluss durch eine bestimmte Reaktion in der Reaktionsrate variiert. Somit kann beispielsweise der Effekt von einer abnehmenden Expressionsrate

eines bestimmten Enzyms beziehungsweise des Gens auf die Wachstumsrate untersucht werden.

- **Simulieren von Gen-Knockout-Phänotypen:**
Die Auswirkungen eines Gen-Knockouts auf die Zellwachstumsrate kann auch durch die Lineare Optimierung vorhergesagt werden. Hierfür werden die obere und untere Grenze der Reaktionsrate des Enzyms, dessen Gen ausgeschaltet werden soll, auf Null gesetzt. Anschließend kann ausgewählt werden, ob durch eine FBA oder über den MOMA-Algorithmus (*Minimization Of Metabolic Adjustment*) der Effekt berechnet werden soll. Während bei der FBA einfach das maximale Wachstum erneut berechnet wird (ohne die spezifische Reaktion), versucht MOMA [109] die Abweichungen der metabolischen Flüsse zwischen der optimalen Flussverteilung des Wildtyps und des Gen-Knockouts minimal zu halten.
- **Flussvariabilitätsanalyse:**
Mithilfe der COBRA-Toolbox ist es möglich, die Variabilität eines Flusses in Bezug auf die Maximierung der Zielfunktion zu berechnen. Dies bedeutet, wie stark der Fluss durch eine bestimmte Reaktion variieren darf unter Berücksichtigung, dass das Maximum der Zielfunktion um einen bestimmten prozentualen Wert abweichen darf. Natürlich ist es möglich, keine Abweichung zu erlauben und trotzdem die Variabilität jeder einzelnen Reaktion zu ermitteln.

Nachteile der Flussbilanzanalyse

Selbstverständlich besitzt auch die Anwendung FBA ihre Grenzen und Einschränkungen. Eines der wichtigsten Kriterien der FBA ist die Zielfunktion, welche für die spezifische Zelle beziehungsweise den spezifischen Zelltyp im Vorfeld bekannt sein muss. Hier ist leicht zu erkennen, dass diese Zielfunktion vor allem für multizelluläre Organismen nicht leicht zu definieren ist. Allgemein ist es nicht möglich, dass eine einzelne Zielfunktion den Zustand der metabolischen Flüsse für jede mögliche Wachstumsbedingung beschreibt [110].

Ein weiteres Defizit der FBA besteht darin, dass durch das Lösen des Optimierungsproblems, nur eine einzige Flussverteilung geliefert wird. Falls mehrere optimale Lösungen existieren, welches aufgrund alternativer Stoffwechselwege sehr oft der Fall ist, gibt es keine Garantie, dass diese bestimmte Flussverteilung der gesuchten *in vivo* Flussverteilung entspricht.

2.1.3. Rekonstruktion metabolischer Netzwerke

Die Rekonstruktion metabolischer Netzwerke folgt meist einem Standardverfahren mit vier Schritten, welches im Detail in der Literatur beschrieben wurde [48, 111, 112]. Als erster Ausgangspunkt dient die annotierte Genomsequenz des spezifischen Organismus aus der GenBank-Datenbank (<http://www.ncbi.nlm.nih.gov/genbank/>) [113] beziehungsweise aus einer spezifischen Datenbank, wie zum Beispiel für Cyanobakterien die CyanoBase-Datenbank (<http://genome.microbedb.jp/cyanobase/>) [114, 115],

die zusammen mit einer Stoffwechseldatenbank, wie unter anderem der *Kyoto Encyclopedia of Genes and Genomes* (KEGG; <http://www.genome.jp/kegg/>) [51], einen ersten Entwurf des Netzwerkes für den spezifizierten Organismus ergeben.

Als nächstes wird versucht, Lücken und Unstimmigkeiten innerhalb dieses ersten Entwurfs des Netzwerkes zu identifizieren und diese manuell auf Grundlage und unter Zuhilfenahme der verfügbaren biochemischen Literatur, zu kuratieren. Jedoch muss beachtet werden, dass nicht alle identifizierten, fehlenden Schritte auch fehlerhaften Lücken im jeweiligen metabolischen Netzwerk entsprechen. Besonders Cyanobakterien sind für ihr sehr kleines und „schlankes“ Genom bekannt, wie vor allem die Gattung *Prochlorococcus* [36]. Bei einem anderen Beispiel, dem marinen Cyanobakterium UCYN-A, wurde erst vor kurzem gezeigt, dass der Sauerstoff erzeugende Photosystem II-Komplex (PS II) in der Elektronentransportkette fehlt. Stattdessen zeigt UCYN-A einen stark eingeschränkten, photofermentativen Stoffwechsel und ist wahrscheinlich abhängig von essentiellen Verbindungen, die von anderen Organismen synthetisiert werden [116]. In einem solchen Fall würde eine Vervollständigung dieser Stoffwechsellücke zu fehlerhaften Lösungen führen und Stoffwechselwege einbeziehen, die nicht im jeweiligen Organismus vorhanden sind.

Sobald innerhalb der Rekonstruktion die Synthese von allen relevanten Stoffwechselzwischenprodukten ermöglicht ist, wird der gesamte Satz von Reaktionen, also das metabolische Netzwerk, in ein mathematisches Modell transferiert. Bei dieser Umwandlung in ein Rechenmodell werden mehrere zusätzliche Reaktionen, sogenannte Hilfsreaktionen in das Netzwerk integriert, welche nicht zu den „echten“ enzymatischen Reaktionen zählen. Zu diesen Hilfsreaktionen gehört unter anderem die Bildung von Biomasse, welches einem zellulären Wachstum entspricht (siehe Kapitel 2.1.9). Eine weitere wichtige Hilfsreaktion ist die Reaktion für die zelluläre *Maintenance*, ein basaler Energieverbrauch, der für die Erhaltung zusätzlicher, nicht modellierter zellulärer Prozesse benötigt wird. Dieser basale Energieverbrauch wird in der Regel durch das Hinzufügen einer ATP-verbrauchenden Reaktion (ATP-Hydrolyse) simuliert, aber auch die Nutzung von NADPH oder ein genereller Metabolitverbrauch können ebenfalls in dieser Funktion berücksichtigt werden. In den metabolischen Modellen in dieser Arbeit wird der Energieverbrauch durch einen basalen ATP-Verbrauch repräsentiert.

Darauffolgend muss das Netzwerk noch in ein computerlesbares Format umgewandelt werden. Dabei wird auf eine einheitliche Annotation sowie eine leichte Austauschmöglichkeit des Formates geachtet. Geeignet ist zum Beispiel das etablierte SBML-Format (*Systems Biology Markup Language*) [117]. Das SBML-Format wird als Standard-Austauschformat für die stöchiometrische Analyse von mathematischen Modellen gesehen. Zu diesem Zweck stehen mehrere Software-Pakete zur Verfügung, welche in der Lage sind, das SBML-Format zu lesen [104–107]. Abschließend wird das Modell unter Verwendung von *Constraint*-basierten und anderen Methoden der Systembiologie analysiert, mit denen es insbesondere möglich ist, Vorhersagen über optimale Flussverteilungen zu treffen, um zum Beispiel phänotypische Daten zu validieren.

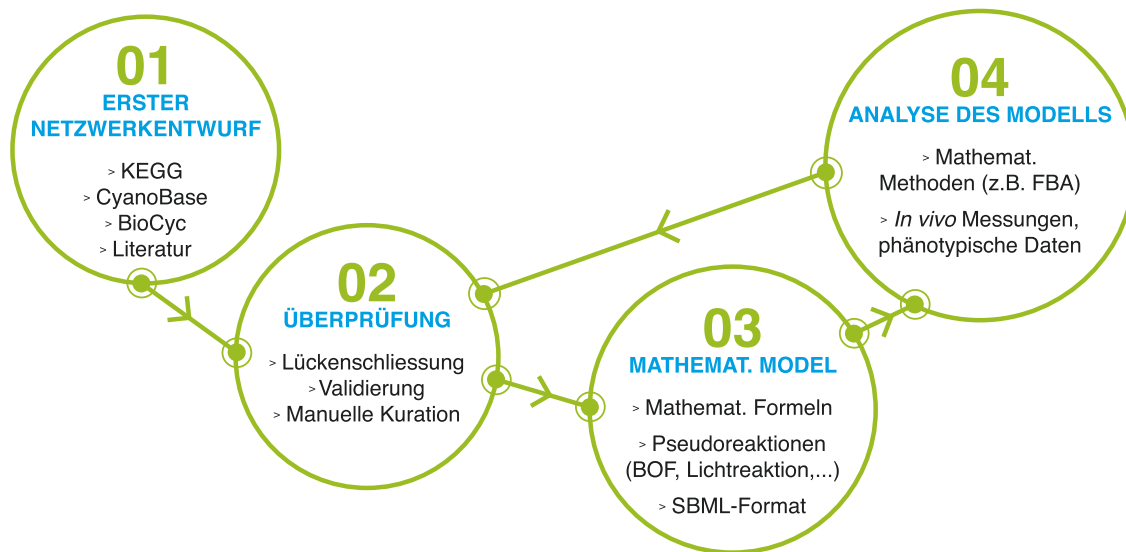


Abbildung 2.2.: Schematischer Ablauf einer Netzwerkrekonstruktion in vier Schritten: Erster Entwurf des Netzwerkes, Überprüfung und Kuration des Netzwerkes unter Zuhilfenahme von Fachliteratur und die Übertragung in ein mathematisches Modell, welches mithilfe von mathematischen Methoden analysiert und mit experimentell gewonnenen Daten verglichen wird. Die gewonnenen Informationen können iterativ in die Überprüfung des Netzwerkes mit einfließen.

Der gesamte Prozess ist iterativ und es sollte daher bedacht werden, dass nach diesen Schritten und vollständiger Überprüfung des Modells, keine fehlerfreie Darstellung des metabolischen Netzwerkes des jeweiligen Organismus zu erwarten ist. Vielmehr stellt es einen Ausgangspunkt für eine systematische funktionale Verifizierung von nicht identifizierbaren oder unbekannten enzymatischen Stoffwechselschritten, welche wiederum zu weiteren Experimenten führen, dar. Der gesamte Prozess wird anhand eines Schemas in Abbildung 2.2 verdeutlicht.

2.1.4. Rekonstruktion des metabolischen Netzwerkes von *Synechocystis* sp. PCC 6803

Das Ziel einer metabolischen Netzwerkrekonstruktion ist eine umfassende Einbeziehung aller biochemischen Umwandlungen, einschließlich Transport- und nichtenzymatischen Reaktionen, die innerhalb einer Zelle oder eines Organismus ablaufen. Als Ausgangspunkt für die Rekonstruktion des metabolischen Netzwerkes von *Synechocystis* diente die Cyanobakterien-Genomdatenbank CyanoBase [114, 115] mit der annotierten Sequenz des Genoms von *Synechocystis* [13].

Mithilfe der Genannotation konnte eine anfängliche Liste von metabolischen Enzymen erstellt werden. Diesen Enzymen wurden anschließend die zugehörigen metabolischen Reaktionen unter Verwendung der KEGG- und der Biocyc-Datenbank (<http://www.biocyc.com/>) [52] zugeordnet. Durch die Kombination der beiden Ressourcen konnte ein erster Entwurf des Netzwerkes, welcher aus Gen-Protein-Reaktion-As-

soziationen für jedes annotierte Gen und einer Reihe von zusätzlichen nicht enzymatischen (spontanen) Reaktionen und Transportprozessen bestand, erstellt werden. Darauf folgend wurden Ansätze unternommen, die bestehenden Lücken im Netzwerk iterativ zu schließen. Zu diesem Zweck wurden zunächst unvollständige Stoffwechselwege identifiziert, indem überprüft wurde, ob das jeweilige Produkt oder der jeweilige Endmetabolit unter phototrophen Bedingungen synthetisiert werden kann. Besonders alle Stoffe, die in der Biomassereaktion aufgeführt werden, müssen innerhalb des Netzwerkes synthetisierbar sein. Zudem müssen Metabolitbestände jeder internen Spezies, die während des Wachstums vermindert werden, durch das Modell wieder auffüllbar sein. Um die identifizierten Lücken in den Stoffwechselwegen zu schließen, wurden verschiedene Stoffwechsel-Datenbanken, wie die KEGG-, die BioCyc- und die BRENDA-Datenbank (*Braunschweig ENzyme DAtabase*; <http://www.brenda-enzymes.org/>) [50] sowie biochemische Fachliteratur durchsucht. Ein konkretes Beispiel für die Notwendigkeit der Einbeziehung spezieller Literatur zeigte sich bei *Synechocystis* für den Fall der Photorespiration beziehungsweise der dazugehörigen Stoffwechselwege. Die Aufklärung der einzelnen Schritte der Photorespiration, war zum Zeitpunkt der Rekonstruktion [93] für *Synechocystis* in den genannten Stoffwechseldatenbanken nicht beschrieben und konnte nur im Detail aus der Literatur von Eisenhut et al. [118, 119] entnommen werden.

Wenn die Literatursuche nicht erfolgreich war, wurde recherchiert, welche Reaktionen, in anderen Organismen vorhanden sind, um die Synthese des gewünschten Produktes zu gewährleisten. Hierbei begann die Suche zuerst in anderen Cyanobakterienstämmen und anschließend wurde diese auf weiter entfernte Organismen wie andere Bakterien oder Pflanzen ausgedehnt. Unter Verwendung des BLAST-Algorithmus (http://blast.kazusa.or.jp/blast_search/cyanobase/Gene) [55] wurden zutreffende Kandidaten-Gene im Genom von *Synechocystis* identifiziert. Ein Beispiel für diesen Vorgang ist das Enzym Methylentetrahydrofolatreduktase (EC 1.5.1.20), welches für *Synechocystis* nicht annotiert ist, aber ein entsprechendes Gen für *Cyanothece* PCC 8801 gefunden werden kann. Diese Erkenntnis ermöglichte die Identifizierung eines Kandidaten-Gens slr2141 für dieses Enzym im Genom von *Synechocystis* sp. PCC 6803.

Wenn eine solche Suche durch einen direkten Vergleich kein zufriedenstellendes Ergebnis lieferte, wurde in der primären Literatur nach einer Beschreibung des jeweiligen Stoffwechselweges recherchiert, beginnend mit dem Stoffwechsel von Cyanobakterien und anschließender Ausdehnung der Suche auf Pflanzen und andere Bakterien und schließlich auf alle restlichen Organismen. Eine Liste aller neu annotierten Gene, die für metabolische Enzyme kodieren könnten, sind in der Tabelle A.2 im Anhang aufgeführt. Auch Details zum Vergleich, wie die jeweiligen Scores und E-Werte, sind in dieser Liste zu finden.

Im Laufe des Rekonstruktionsprozesses fiel auf, dass die Spezifität für den jeweiligen Cofaktor für viele Enzyme bezüglich NAD(H) und NADP(H) unklar beziehungsweise nicht bekannt ist. Da das metabolische Netzwerk von *Synechocystis* über eine NAD-NADP-Transhydrogenase (slr1239, slr1434) verfügt, die eine direkte Konvertierung zwischen NADH und NADPH in beide Richtungen ermöglicht, wurden für solche

Fälle generell nur NADP(H) als Cofaktor für die jeweilige Reaktion ins Modell übernommen.

2.1.5. Annahmen für das metabolischen Netzwerk von *Synechocystis* sp. PCC 6803

Obwohl mehrere Rekonstruktionen des Cyanobakteriums *Synechocystis* schon seit längerem in der Literatur verfügbar sind [120–126], wurden nur wenige Versuche unternommen, um das fehlende metabolische Wissen zu systematisieren. Tatsächlich sind mehrere Aspekte des metabolischen Netzwerkes und seiner wichtigsten Synthesewege noch unzureichend erforscht. Zum Beispiel fehlen innerhalb der aktuellen Rekonstruktion [94] komplette Synthesewege für die Aminosäuren Methionin und Asparagin. Ferner sind auch die Synthesewege für Serin, Glycin und Alanin wahrscheinlich unvollständig beziehungsweise nicht vollständig aufgelöst. Um trotzdem die Synthese von Methionin und Asparagin für die Biomasse zu ermöglichen, wurden die Reaktionsschritte für Methionin (EC 2.3.1.31, EC 2.5.1.49) von dem Cyanobakterium *Microcystis aeruginosa* NIES-843 übernommen. Für Asparagin wird angenommen, dass eine Asparagin-Synthetase (EC 6.3.5.4) die Umsetzung von Asparagin aus Aspartat katalysiert. Für alle anderen proteinogenen Aminosäuren sind die Synthesewege und deren vermeintliche enzymatische Schritte annotiert oder konnten im Laufe der Rekonstruktion aufgeschlüsselt werden.

Um längere Perioden ohne Licht zu überleben, nutzen Cyanobakterien Glykogen, Cyanophycin und Poly- β -Hydroxybutyrat (PHB) als natürliche interne Speicherverbindungen. Nichtsdestoweniger gilt auch für die Speicherstoffe, dass deren Stoffwechselwege, also der Aufbau und die spätere Mobilisierung, nicht komplett aufgelöst sind. Während der Syntheseweg von internem PHB vollständig aufgeklärt ist und bereits in der Literatur beschrieben wurde [127], sind die notwendigen Schritte zum enzymatischen Abbau bei Cyanobakterien zum jetzigen Zeitpunkt nicht bekannt. Zudem sind enzymatische Schritte zur Synthese mehrerer Zellwandkomponenten, wie UDP-Glucose (UDP-Glucose-Phosphorylase; EC 2.7.7.9) und vor allem Glycerolipide lückenhaft beziehungsweise nicht annotiert. Auch die Synthesewege der Cofaktoren Vitamin B₆ und B₁₂ weisen einige fehlende enzymatische Schritte auf.

Eine wesentliche Unklarheit für *Synechocystis* betrifft den Elektronentransport innerhalb der Elektronentransportkette. Es ist nicht vollständig geklärt, ob das pflanzliche Plastochinon oder das bakterielle Ubichinon als Elektronentransporter fungiert. Während der pflanzliche Syntheseweg des Plastochinons nahezu vollständig in *Synechocystis* annotiert ist, zeigen Knockout-Studien von Dähnhardt et al. (2002) [128], dass eine Ausschaltung dieses Stoffwechselweges keinen Effekt auf die photosynthetischen Funktionen von *Synechocystis* hat. Daher könnte es sein, dass ein alternativer Stoffwechselweg für die Synthese von Plastochinon in Cyanobakterien beziehungsweise *Synechocystis* existiert, oder *Synechocystis* stattdessen Ubichinon verwendet. Ein alternativer Weg für die Synthese von Plastochinon bei Cyanobakterien wurde erst kürzlich von Sadre et al. (2012) [129] vorgeschlagen.

Weitere fehlende enzymatische Schritte wurden durch gezielte BLAST-Suchen identifiziert und geschlossen. Hierfür wurde die Suche zunächst mit nah verwandten Cyanobakterien begonnen [130]. Darüber hinaus wurde extensiv in der verfügbaren biochemischen Literatur recherchiert, um mögliche alternative enzymatische Wege zu identifizieren. Eine Liste der fehlenden oder unklaren enzymatischen Schritte und notwendigen Ergänzungen, damit alle Bedingungen für die Synthese der Biomasse und damit das Wachstum *in silico* gewährleistet wird, sind in der Tabelle A.1 im Anhang aufgeführt.

2.1.6. Die Synthese von Glycin und Serin

Für Serin und Glycin wurde angenommen, dass eine Synthese über den Glyoxylat-Stoffwechsel möglich ist, welcher durch Eisenhut et al. (2008) [119] detailliert beschrieben wurde. Dabei dient insbesondere Glyoxylat, welches während der Photorespiration zwangsläufig gebildet wird, als Vorstufe für die Aminosäuren Glycin, Serin und Cystein.

Eine weitere Möglichkeit der Serin-Synthese besteht durch die Existenz eines Phosphoserin-Weges. In diesen Weg sind eine 3-Phosphoglycerat-Dehydrogenase (EC 1.1.1.95), eine Phosphoserin-Transaminase (EC 2.6.1.52) und eine Phosphoserin-Phosphatase (EC 3.1.3.3) involviert. Von diesen drei Enzymen ist alleinig die 3-Phosphoglycerat-Dehydrogenase zum Zeitpunkt der Rekonstruktion in den Stoffwechsel- und Genomdatenbanken für *Synechocystis* annotiert. Es konnte jedoch gezeigt werden, dass für *Hydrogenobacter thermophilus* eine Phosphoglycerat-Mutase die Funktion einer Phosphoserin-Phosphatase übernehmen kann [131]. Da mehrere dieser Phosphoglycerat-Mutasen im Genom von *Synechocystis* auffindbar sind [130], ist es sehr wahrscheinlich, dass auch für *Synechocystis* eine Phosphoserin-Phosphatase-Aktivität einer Phosphoglycerat-Mutase existiert.

Aufgrund dieser Hypothese wurden experimentelle Ansätze für ein solches Szenario definiert, mit denen eine Aktivität der Reaktion der Phosphoserin-Phosphatase in Zellextrakten von *Synechocystis* nachgewiesen werden konnte. Anschließend war es möglich, die zugehörige Phosphoglycerat-Mutase zu identifizieren (slr1124; persönliche Kommunikation Wolfgang Lockau). Ferner konnte im Labor die Aktivität des gesamten Phosphoserin-Weges in *Synechocystis* nachgewiesen und die entsprechende Transaminase identifiziert werden (persönliche Kommunikation Wolfgang Lockau und Martin Hagemann). Abschließend wurde daraufhin der komplette Phosphoserin-Weg in das stöchiometrische Modell von *Synechocystis* übernommen.

2.1.7. Formale Eigenschaften des metabolischen Netzwerkmodells

Eine der wichtigsten Eigenschaften eines Netzwerkes ist, dass das Netzwerk in sich konsistent ist und alle Ladungen und Atome, besonders auf der Ebene der Protonen, innerhalb der Reaktionen ausgeglichen sind. Für das *Synechocystis*-Netzwerk wurden diese Kriterien unter Verwendung der COBRA-Toolbox [105] und der SuBliMiNaL-Toolbox (durchgeführt von Neil Swainston) [132] überprüft und gegebenenfalls kor-

rigiert. Für den Protonen- und Ladungsausgleich ist der vorherrschende pH-Wert des Milieus entscheidend. Unter Standardbedingungen variiert der intrazelluläre pH-Wert von *Synechocystis* im Laufe des Tages zwischen ca. 7,2 und 7,8 [133]. Der Einfachheit halber und aufgrund der Datenlage wurde bei der Rekonstruktion ein pH-Wert von 7,00 als Referenzzustand angenommen.

Zur Durchführung der Simulationen und dem Transfer des Modells in verschiedene Analyse-Softwares wurde das metabolische Modell im MIRIAM-kompatiblen SBML-Standard [134] exportiert. Unter anderem wurden dabei alle Metaboliten des Netzwerkes, wenn möglich mit ihrer entsprechenden CheBI ID (*Chemical Entities of Biological Interest*) [135], referenziert. Zudem wurde eine SBML-Version des Netzwerkes speziell für die Analysen unter der Verwendung der COBRA-Toolbox erstellt.

2.1.8. Kriterien für die Modellierung des metabolischen Netzwerkes

Für die Flussbilanzanalyse des phototrophen Wachstums von *Synechocystis* sp. PCC 6803 wurde angenommen, dass ausschließlich Nitrat als einzige Stickstoffquelle fungiert und über einen ATP-abhängigen Transport in die Zelle gelangt. Eine Verwendung von Harnstoff oder Ammoniak, welche ebenfalls als Stickstoffquelle für *Synechocystis* fungieren könnten, wurde für die Simulationen nicht berücksichtigt. Die Aufnahme von Kohlenstoff wurde ausschließlich in Form von Hydrogencarbonat über einen ATP-abhängigen Transporter [136] beschränkt. Das heterotrophe beziehungsweise mixotrophe Wachstum, welches mit der Aufnahme von Verbindungen mit mehreren Kohlenstoffatomen wie Glucose einhergeht, wurde nicht genauer untersucht. Auch wurde die Aufnahme von komplexeren Verbindungen, wie zum Beispiel von Aminosäuren oder Putrescin, nicht in den Simulationen betrachtet.

Für die Simulationen wurde nur der Licht-Eingang in Form von der Anzahl einfallender Photonen als Austauschfluss festgelegt, während die Aufnahme von Nitrat, Sulfat und Kohlenstoff (in Form von Hydrogencarbonat) nicht beschränkt wurde. Durch diese Vorgehensweise wurde die Optimierung über zwei Schritte, wie es bei Shastri und Morgan (2005) [120] der Fall ist, auf einen Schritt reduziert. Diese hatten zuerst die Bildung von Biomasse mit einer festen Kohlenstoffaufnahme und unbegrenztem Licht optimiert und anschließend den Licht-Eingang bei gleichbleibender Biomassebildung als zusätzliche Nebenbedingung minimiert. Für die Simulationen des phototrophen Wachstums mittels der Flussbilanzanalyse wurde immer von einer Verdopplungszeit von 24 Stunden für *Synechocystis* sp. PCC 6803 ausgegangen, welches einer Wachstumsrate von ungefähr $0,0289 \text{ h}^{-1}$ entspricht.

Neben der BOF wurde zusätzlich ein basaler ATP-Verbrauch (*Maintenance*) für jede Flussoptimierung angenommen. Diese entspricht einem allgemeinen ATP-Verbrauch aller übrigen zellulären Prozesse, die nicht vom metabolischen Netzwerk abgedeckt werden, wie zum Beispiel Regulationsprozesse. Da keine experimentellen, zuverlässigen Daten in der Literatur zu dieser Thematik zur Verfügung standen, wurde der Wert auf ca. 10 % der maximalen ATP-Erzeugung durch die Photosynthese festgelegt. Es stellte sich heraus, dass die basale ATP-Nutzung, im Gegensatz zu der Situation in heterotrophen Organismen, keinen qualitativen Einfluss auf die FBA-Lösung besitzt, da

Tabelle 2.1.: Eingangsparameter für die FBA des metabolischen Netzwerkes von *Synechocystis* sp. PCC 6803. Aufgeführt sind die festgelegten Flussraten [$\text{mmol gDW}^{-1} \text{h}^{-1}$] für die entsprechenden Reaktionen im Netzwerk für das phototrophe Wachstum (Licht) und bei Abwesenheit von Licht (Dunkel).

Reaktion	Reaktionsnummer	Licht	Dunkel
Lichteingang	R443	18,7	0
Diverse Kohlenstoff-transporter	R674	0	0
	R675	0	0
	R676	0	0
	R677	0	0
	R678	0	0
	R679	0	0
	R680	0	0
	R681	0	0
	R682	0	0
	R685	0	0
	R686	0	0
	R687	0	0
	R712	0	0
	R713	0	0
basaler ATP-Verbrauch	R147	0,13	0,07
Cytochrom-c-Oxidase Mehler-ähnliche Reaktion	R457	0,2263	0
	R481	0,2263	0
Sauerstoffradikale am PS II Mehler-Reaktion	R480	0,0477	0
	R482	0,0473	0
Glykogenmobilisierung	R765	0	0,01
Biomassezielfunktion	R147	max.	max.

durch eine simple Erhöhung der Lichtintensität beziehungsweise der Photonenanzahl dieser Wert ausgeglichen werden kann.

Für die Simulation von Wachstum bei Abwesenheit von Licht wurde einzig Glykogen als Speicherstoffverbindung berücksichtigt. Ferner wird für dieses Szenario festgelegt, dass die Reaktion der Glutamat-Dehydrogenase (EC 1.4.1.4) (GDH) irreversibel abläuft. Andernfalls wäre es möglich, dass diese spezifische Reaktion, zusammen mit anderen Reaktionen, als zusätzliche Transhydrogenase fungieren könnte, was es für manche Simulationsszenarien bei Abwesenheit von Licht, in denen eine inaktive Transhydrogenase erwünscht war, zu vermeiden galt. Alle Eingangsparameter der FBA des *Synechocystis*-Modells für das phototrophe Wachstum sowie das Wachstum bei Abwesenheit von Licht, sind in der Tabelle 2.1 aufgeführt.

Die Optimierung der Biomasseproduktion ergibt in den meisten Fällen keine eindeutige Lösung, sondern es ist möglich, dass viele Flusslösungen die Optimierungskriterien erfüllen. Um hier weiter anzusetzen, können noch weitere Einschränkungen bei der Optimierung als Kriterium dienen. Es ist zum Beispiel möglich, auch die Minimierung der gesamten Flusswerte als weitere Bedingung für das Optimierungsproblem festzulegen. Dies wird normalerweise angewandt, um die Minimierung der Synthesekosten von metabolischen Enzymen, die für eine bestimmte Flusslösung benötigt werden, zu repräsentieren [137]. Alle für diese Arbeit relevanten Flussverteilungen wurden unter Berücksichtigung dieser zusätzlichen Bedingung berechnet.

2.1.9. Die Biomassezielfunktion

Eine besondere Bedeutung für die FBA hat die Definition einer Biomassezielfunktion (*Biomass Objective Function*) (BOF), die meistens als Zielfunktion bei einer FBA fungiert. Diese Biomassefunktion ist eine Hilfsreaktion, welche das zelluläre Wachstum repräsentiert. Sie spezifiziert die molaren Verhältnisse aller zellulären Komponenten beziehungsweise metabolischen Vorstufen, die nötig sind, ein Gramm Trockengewicht von kompletten Zellen zu synthetisieren. Die Biomassezielfunktion kann sich je nach den experimentellen Bedingungen ändern. Bei Cyanobakterien wird deswegen oft zwischen heterotrophem, mixotrophem oder phototrophem Wachstum unterschieden, da sich bei diesen Zuständen auch die zelluläre Zusammensetzung ändert [120]. Für diese Arbeit wurde nur eine Biomassezusammensetzung für phototrophes Wachstum berücksichtigt.

Eine erste Definition der Biomassefunktion für *Synechocystis* wurde schon 2005 von Shastri und Morgan [120] und 2008 von Fu [122] vorgeschlagen. Allerdings basierte diese Biomassefunktion vor allem auf Vorstufen von Aminosäuren als Bestandteil der Biomassezusammensetzung und war somit sehr beschränkt. Eine deutlich aufwändigere Funktion wurde durch Cogne et al. (2003) [138] für ein metabolisches Modell von *Arthrospira platensis*, auch bekannt unter *Spirulina platensis*, einem weiteren Cyanobakterium, bereitgestellt. Diese Funktion setzte sich aus fünf Komponenten von Makromolekülen zusammen: Kohlenhydrate, Proteine, Lipide, RNA und DNA und das Pigment Chlorophyll. Eine modifizierte Version dieser Biomasseformel wurde in den ersten Versionen des metabolischen Netzwerkes von *Synechocystis* verwendet [93].

Für die aktuelle Version des Netzwerkes wurde die Biomassefunktion, welche in Nogales et al. (2012) [125] aufgeführt wird, modifiziert und verfeinert. Zu diesem Zweck wurden zuerst die verschiedenen Bestandteile der Biomasse in sieben verschiedene Biomassekomponenten eingeteilt und einzelne Makromoleküle definiert: Proteine, DNA, RNA, Zellwandbestandteile, Lipide, lösliche Zellkomponenten, anorganische Ionen und Pigmente. Außerdem wurden zusätzliche Metaboliten, die bei den Metabolitbestimmungen in den Zusatzinformationen von Nogales et al. (2012) [125] aufgeführt wurden, in die Biomasseformel integriert. Dazu gehört unter anderem Vitamin B₆. Überdies wurde ein wachstumsabhängiger ATP-Verbrauch für den Energiebedarf von Proteinsynthesen, Regulation und Wachstum mit einbezogen. Die komplette Biomasseformel ist in Tabelle 2.2 aufgeführt.

Tabelle 2.2.: Die in dieser Arbeit verwendete Biomasseformel des metabolischen Netzwerkes von *Synechocystis* sp. PCC 6803. Die Biomasse ist dabei aufgeteilt in die Komponenten: Proteine, DNA, RNA, Lipide, Zellwand, lösliche Zellkomponenten, anorganische Ionen und Pigmente. Sie enthält zusätzlich Glykogen als Speicherverbindung und einen wachstumsabhängigen ATP-Verbrauch (Tabelle aus Knoop et al. (2013) [94]).

Zusammensetzung	
Proteine	0,766 L-Alanin + 0,455 L-Arginin + 0,366 L-Asparagin + 0,454 L-Aspartat + 0,090 L-Cystein + 0,502 L-Glutamin + 0,546 L-Glutamat + 0,666 Glycin + 0,168 L-Histidin + 0,566 L-Isoleucin + 1,029 L-Leucin + 0,378 L-Lysin + 0,177 L-Methionin + 0,362 L-Phenylalanin + 0,463 L-Prolin + 0,524 L-Serin + 0,496 L-Threonin + 0,140 L-Tryptophan + 0,263 L-Tyrosin + 0,608 L-Valin
DNA	0,7749 dCTP + 0,8517 dTTP + 0,7772 dGTP + 0,8487 dATP
RNA	0,8080 GTP + 0,7300 ATP + 0,7799 CTP + 0,8012 UTP
Zellwand	0,4276 Lipid A-Disaccharid + 0,4276 Peptidoglycan
Lipide	0,380 3- β -D-Galactosyl-1,2-Diacylglycerol + 0,2650 Digalactosyl-Diacylglycerol + 0,2550 Sulfoquinovosyldiacylglycerol + 0,100 Phosphatidylglycerol
lös. Zellkomponenten	1,147 Putrescin + 0,233 Spermidin + 0,010 Acetyl-CoA + 0,006 CoA + 0,003 Succinyl-CoA + 0,001 Malonyl-CoA + 0,062 NAD ⁺ + 0,002 NADH + 0,004 NADPH + 0,012 NADP ⁺ + 0,008 FAD ⁺ + 0,008 Tetrahydrofolat + 0,008 5,10-Methenyltetrahydrofolat + 0,008 5-Methyltetrahydrofolat + 0,008 Thiamin-Diphosphat + 0,008 Pyridoxalphosphat + 0,008 Heme O + 0,008 Heme A + 0,008 Heme + 0,008 Glutathion + 0,008 Cobalamin + 0,002 Undecaprenyl-Diphosphat + 0,008 10-Formyltetrahydrofolat + 0,008 Chorismat + 0,008 S-Adenosylmethionin + 0,008 Riboflavin
anorgan. Ionen	17,040 K ⁺ + 1,136 NH ₄ ⁺ + 0,757 Mg ²⁺ + 0,454 Ca ²⁺ + 0,682 Fe ²⁺ + 0,682 Fe ³⁺ + 0,303 Cu ²⁺ + 0,303 Mn + 0,303 H ₂ MoO ₄ + 0,303 Co ²⁺ + 0,303 Zn ₃ Cu ²⁺ + 0,379 Sulfat + 0,379 Orthophosphat + 0,379 Na ⁺
Pigmente	0,827 Chlorophyll α + 0,134 β -Carotin + 0,113 Zeaxanthin + 0,098 Echinenon + 0,063 γ -Carotin + 0,010 α -Tocopherol + 0,001 β -Tocopherol + 0,001 γ -Tocopherol + 0,009 δ -Tocopherol + 0,067 Phyllochinon
Biomasse	53,350 ATP + 53,350 H ₂ O + 0,51 Proteine + 0,031 DNA + 0,17 RNA + 0,059 Zellwand + 0,12 Lipide + 0,029 lös. Zellkomponenten + 0,01 anorgan. Ionen + 0,0244 Pigmente + 0,21031 Glykogen

2.1.10. Zeitaufgelöste Flussbilanzanalyse

Für die Modellierung eines vollständigen Tagesablaufes mithilfe der Flussbilanzanalyse wurde zunächst von einer Verdopplungszeit von etwa 24 Stunden ausgegangen. Da bei der FBA der Fluss durch die einzelnen Biomassekomponenten in mmol pro Stunde pro Gramm Trockengewicht gegeben ist, wurden die Flusswerte über 24 Stunden aufintegriert, um eine Schätzung der Gesamtmenge jedes Bestandteils über den Tagesverlauf zu erhalten. Basierend auf diesen angegebenen Mengen, wurde ein artifizielles Szenario für die Simulation eines kompletten Tagesdurchlaufs postuliert:

- Eine kontinuierliche Synthese der DNA- und der anorganischen Ionen-Biomassekomponenten.
- Die Synthese der Protein-, RNA-, Zellwand- und Lipid-Komponenten wurde auf die Lichtphase beschränkt.

- Für die Synthese der Pigmente wurden die Parameter für die Simulation so gewählt, dass diese zwei Stunden vor dem Beginn der Lichtphase startet und zwei Stunden vor dem Beginn der Dunkelphase endet.
- Um den Bedarf an Energie in der Dunkelphase zu erfüllen, wenn kein Licht zu- gegen ist um die beiden Photosysteme anzuregen, wird angenommen, dass Gly- kogen, welches ab Mitte der Lichtphase aufgebaut wird, als Speicherverbindung verwendet wird.
- Während der Lichtphase wurde eine stetige Respirationsrate der Cytochrom-c- Oxidase von $0,2263 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ angenommen, dies entspricht ca. 10 % der maximalen photosynthetischen Sauerstoffproduktion des Photosystems II. Wäh- rend der Dunkelphase wird die Respirationsrate auf $0,005 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ redu- ziert.
- Für die Aufrechterhaltung von zellulären Prozessen wird ein basaler ATP-Ver- brauch (*Maintenance*) von $0,13 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ während der Lichtphase festge- setzt. In der Dunkelphase wird der ATP-Verbrauch nicht festgesetzt, aber die Synthese von ATP als Zielfunktion definiert.
- In der Simulation wurde für den Lichteinfall ein einfaches Szenario in einer dreieckigen Form gewählt, beginnend mit einem linearen Anstieg bei der circa- dianen Zeit $CT = 0 \text{ h}$ und einem Maximum bei $CT = 6 \text{ h}$ mit einem Wert von $9,44 \text{ mmol Photonen gDW}^{-1} \text{ h}^{-1}$ und einem anschließenden linearen Abfall.

Tabelle 2.3.: Modifikationen der Faktoren für die Biomassekomponenten Proteine, DNA, RNA, Lipide, Zellwand, lösliche Zellkomponenten, anorganische Ionen und Pigmente zu bestimmten Zeitpunkten des Tages (Hauptschritte) für die zeitabhängige FBA. Die Synthese der DNA- und der anorganischen Ionen- Komponente bleibt über den gesamten Tagesablauf konstant. Die Pigmentsynthese erreicht ihr Maximum am Anfang des Tages, während die Glykogensynthese erst nach $CT = 6 \text{ h}$ beginnt. In den Vormittags- stunden ($CT = 0-6 \text{ h}$) und während der Dunkelphase ($CT = 12-24 \text{ h}$) wurden die Faktoren dem nächstem Hauptschritt angenähert (Tabelle aus Knoop et al. (2013) [94]).

	CT [h]						
	0	6	7	8	9	10	11
Proteine	0,55	0,38	0,39	0,39	0,43	0,46	0,57
DNA	konst.						
RNA	0,18	0,13	0,13	0,13	0,14	0,15	0,19
Zellwand	0,06	0,04	0,04	0,04	0,05	0,05	0,07
Lipide	0,13	0,09	0,09	0,09	0,10	0,11	0,14
lös. Zellk.	0,03	0,02	0,02	0,02	0,02	0,03	0,03
an. Ionen	konst.						
Pigmente	0,04	0,01	0,01	0,01	0,01	0,01	0,00
Glykogen	0,00	0,32	0,32	0,32	0,24	0,19	0,00

Der gesamte Tagesablauf wurde für die Simulation in $24 \cdot 8 = 192$ Optimierungsschritte unterteilt und für jeden einzelnen dieser Schritte wurde eine neue Biomassefunktion festgelegt. Die Übergänge zwischen den unterschiedlichen Biomassezusammensetzungen (Hauptschritten) wurden geglättet, dadurch dass die Differenz der einzelnen Faktoren innerhalb der Biomassefunktion durch die Anzahl der Zwischenschritte dividiert wurde und anschließend für jeden Schritt die Faktoren um diese Werte geändert wurden. Die Hauptschritte der zeitabhängigen Biomassefunktion sind in Tabelle 2.3 gezeigt.

2.2. Ergebnisse und Diskussion

2.2.1. Das metabolische Netzwerk von *Synechocystis* sp. PCC 6803

Im Laufe der Arbeit wurde eine umfassende Rekonstruktion der primären Stoffwechselwege von *Synechocystis* sp. PCC 6803 (*Synechocystis*) entwickelt. Die erste Version dieser Rekonstruktion [93] umfasste die Hauptstoffwechselwege inklusive der photosynthetischen Lichtreaktionen, des Calvin-Benson-Zyklus, der Glykolyse, des Citratzyklus, des Pentosephosphatweges, der oxidativen Phosphorylierung, der Photorespiration, der Synthesen der proteinogenen Aminosäuren sowie der Bereitstellung von Nukleotiden, Lipidvorstufen und Cofaktoren. Dieses Netzwerk wurde unter Berücksichtigung der immer weiter fortschreitenden Kenntnisse über den Organismus sowie aus unterschiedlichen ergänzenden Netzwerkrekonstruktionen neueren Datums [120–126, 139] erweitert.

Das Kernnetzwerk umfasst eine verbundene Sammlung aller bekannten Stoffwechselwege zur Synthese der Biomasse und Cofaktoren, von denen bekannt ist, dass sie im Stoffwechsel von *Synechocystis* vorkommen. So umfasst die aktuelle Version der Netzwerkrekonstruktion von *Synechocystis* [94] zusätzlich zu den vorher genannten Stoffwechselwegen den Lipid- und Fettsäure-Stoffwechsel, die Biosynthese von Peptidoglycan, Chlorophyll *a*, verschiedenen Carotinoiden, Terpenoiden, Chinonen, Tocopherolen und Thiaminpyrophosphat sowie die Synthese von mehreren weiteren Cofaktoren, Vitaminen und vereinzelt stressassoziierten Metaboliten. Zusätzlich wurden im Verlauf der weiteren Beschreibung des Stoffwechsels von *Synechocystis* die photosynthetischen Lichtreaktionen sowie die Transportprozesse qualitativ verbessert und detailreicher beschrieben. Bei der Beschreibung der aufgeführten Stoffwechselwege wurde auf die Vollständigkeit aller Synthesewege in Bezug auf die unterschiedlichen Biomassekomponenten geachtet.

Um profunde Aussagen aus dem Netzwerk zu gewinnen, wurde während dieser Arbeit zwischen einem Kernnetzwerk, welches die vorher genannten, gut validierten Reaktionen enthält und einem erweiterten Netzwerk unterschieden. Das erweiterte Netzwerk beinhaltet, zusätzlich zu den Reaktionen aus dem Kernnetzwerk, noch alle verbleibenden annotierten Enzyme aus dem Genom von *Synechocystis*, welche eine putative Stoffwechselfunktion besitzen (siehe Abbildung 2.3). Dies bedeutet, dass auch unzureichend validierte Reaktionen im erweiterten Netzwerk enthalten sind. Dies führt in vielen Fällen dazu, dass Reaktionen unverbunden sind, ohne dass die jeweilig beteiligten Metaboliten synthetisiert oder weiter verwendet werden können. Aus diesem Grund wurden für die rechnerischen Analysen ausschließlich das Kernnetzwerk verwendet.

Während der Rekonstruktion wurde sichergestellt, dass die Vollständigkeit aller Synthesewege in Bezug auf die Biomassekomponenten, Cofaktoren und Minderung des Metabolitbestandes durch Wachstum gewährleistet wird. Das resultierende metabolische Netzwerk von *Synechocystis* umfasst in der aktuellen Version 677 Gene, die für 491 Enzyme oder Enzymkomplexe kodieren. Hieraus ergeben sich 759 zusammenhängende Stoffwechselreaktionen, bei denen 601 unterschiedliche Spezies umgesetzt

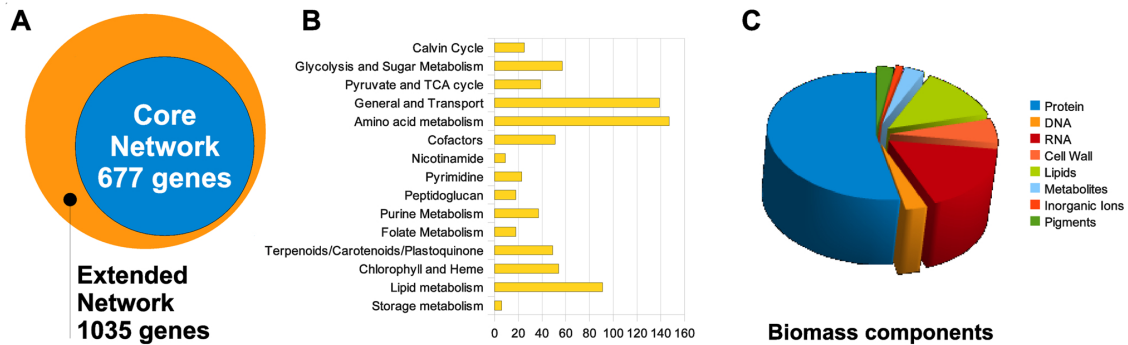


Abbildung 2.3.: Überblick über die Attribute des metabolischen Netzwerkes von *Synechocystis* sp. PCC 6803: A) Unterscheidung zwischen dem Kernnetzwerk und dem erweiterten Netzwerk. B) Verteilung der 759 Reaktionen über die unterschiedlichen Stoffwechselwege des Netzwerkes. C) Aufteilung der unterschiedlichen Komponenten, die für die Biomasse (BOF) benötigt werden (ohne Glykogen und basalem ATP-Verbrauch) (Abbildung aus Knoop et al. (2013) [94]).

werden. Von diesen 759 Reaktionen sind 679 enzymatisch katalysiert, einschließlich 56 Isoreaktionen, sechs spontan ablaufenden beziehungsweise nicht katalysierten Reaktionen, 61 Diffusions- und Transportreaktionen und 13 sogenannten Hilfsreaktionen, wie zum Beispiel Reaktionen der Lichtabsorption, Photophosphorylierung, Biomassebildung oder ein basaler ATP-Verbrauch (*Maintenance*).

Im Modell laufen die Reaktionen innerhalb sieben unterschiedlichen, voneinander getrennten Kompartimenten der Zelle ab. Das sind: Der extrazelluläre Raum, welcher alle Spezies umfasst, die sich außerhalb der Zellmembran befinden, die Zellmembran sowie auch die Thylakoidmembran und der davon eingeschlossene periplasmatische Raum und das Thylakoidlumen. Die meisten metabolischen Reaktionen laufen jedoch im Cytosol der Zelle ab. Zuletzt sind noch die Carboxysomen zu nennen, welche eine wichtige Rolle für die CO₂-Fixierung übernehmen (*CO₂-Concentrating Mechanisms*).

Die Rekonstruktion des metabolischen Netzwerkes von *Synechocystis* konzentriert sich auf eine funktionale Beschreibung des phototrophen Wachstums und die Verwertung von Speicherkomponenten als Energie und Kohlenstoffquelle bei Abwesenheit von Licht. Die Aufnahme und Verwertung von zusätzlichen organischen Kohlenstoffquellen wird bei der mathematischen Modellierung nicht berücksichtigt. Die wichtigsten Eigenschaften des Netzwerkes wurden in Abbildung 2.3 zusammengefasst. Die Rekonstruktion liegt auf der Webseite der elektronischen Version der Veröffentlichung Knoop et al. (2013) [94] im Anhang als SBML- und Excel-Datei vor (<http://www.ploscompbiol.org/article/info%3Adoi%2F10.1371%2Fjournal.pcbi.1003081>).

Zusätzlich wurde eine detaillierte graphische Übersicht des metabolischen Netzwerkes von *Synechocystis* angefertigt, welche im Anhang dieser Arbeit zu finden ist (siehe Abbildung A.1).

2.2.2. Vereinbarkeit bestehender Rekonstruktionen

Während metabolische Rekonstruktionen von phototrophen Organismen im Vergleich zu heterotrophen Mikroorganismen noch unterrepräsentiert sind, wurden vor und während dieser Arbeit einige andere Rekonstruktionen eines metabolischen Netzwerkes von *Synechocystis* sowie von einer Reihe anderer Cyanobakterien publiziert.

Die erste Rekonstruktion von *Synechocystis* wurde bereits 2005 von Shastri und Morgan [120] veröffentlicht und auch unter Verwendung der FBA untersucht. Dieses erste Modell wurde zwei Jahre später von Hong und Lee (2007) [121] erweitert. Beide Rekonstruktionen sind vergleichsweise klein und beschränken sich auf den Primärmetabolismus beziehungsweise den Zentralstoffwechsel von *Synechocystis*, einschließlich der Glykolyse, einem unvollständigen Tricarbonsäurezyklus (*Tricarboxylic Acid Cycle*) (TCA-Zyklus) und dem Calvin-Benson-Zyklus. Zudem beinhalten beide Rekonstruktionen einen Glyoxylatzyklus, dessen Existenz jedoch umstritten ist. Das Vorhandensein des Glyoxylatzyklus in *Synechocystis* wurde zuerst von Yang et al. (2002) [140] vorgeschlagen. Jedoch existieren keine schlüssigen Beweise für diese Behauptung und es wird angenommen, dass dieser im Wesentlichen in die Rekonstruktionen aufgenommen wurde, um einen funktionalen C₂-Zyklus zu ermöglichen (siehe Kapitel 2.2.8).

Eine wesentlich größere genomskalige Rekonstruktion wurde 2008 von Fu [122] veröffentlicht. Diese umfasst etwa 831 Stoffwechselreaktionen. Jedoch enthält diese Rekonstruktion nur eine sehr geringe, manuelle Kuration, was durch spätere Veröffentlichungen bestätigt wurde. Demzufolge existieren sehr viele unvollständige Stoffwechselwege, die aber bei genauerer Betrachtung für eine Modellierung und Simulation unerheblich sind, da bei der Rekonstruktion die eingeschränkte Biomasseformel von Shastri und Morgan [120] (siehe Kapitel 2.1.9) verwendet wurde, wodurch die zusätzlichen Reaktionen keine funktionelle Rolle bei den Simulationen übernehmen. Auch bei der Rekonstruktion von Fu [122] ist der von den Bakterien stammende Glyoxylatzyklus implementiert.

2010 wurde die erste Version des *Synechocystis*-Netzwerkes im Rahmen dieser Arbeit publiziert [93], welche zum Teil noch nicht die hohe Komplexität besaß, wie die aktuelle Version des Netzwerkes. Diese erste Version des Netzwerkes enthielt erstmals eine detaillierte Darstellung der Photorespiration und verzichtete auf eine Einbeziehung des bakteriellen Glyoxylatzyklus. Stattdessen wurde der unvollständige TCA-Zyklus über den alternativen γ -Aminobutyrat (GABA)-Shunt geschlossen und ermöglichte dadurch einen zirkulären Fluss bei der Abwesenheit von Licht. Kurz darauf wurden vier zusätzliche Rekonstruktionen veröffentlicht [123, 124, 139, 141], die jeweils einen unvollständigen TCA-Zyklus und den Glyoxylatzyklus beinhalten.

Erst vor kurzem wurde eine weitere Rekonstruktion von Nogales et al. (2012) [125] vorgestellt, die eine erweiterte Biomassefunktion verwendet, welche aus der Literatur abgeleitet wurde. Diese Rekonstruktion beinhaltet weder einen Glyoxylatzyklus, noch den TCA-Nebenweg, der von Zhang und Bryant [142] für Cyanobakterien beziehungsweise *Synechococcus* sp. PCC 7002 vorgeschlagen wird. Zudem wurden Rekonstruktionen sowohl von *Cyanothece* ATCC 51142 [143] als auch von *Synechococcus* sp. PCC

7002 [144] in letzter Zeit veröffentlicht. Auch ein Vergleich des metabolischen Potentials der beiden Modellorganismen *Synechocystis* sp. PCC 6803 und *Cyanothece* ATCC 51142 wurde vor kurzem publiziert [126].

2.2.3. Flussbilanzanalyse des Netzwerkes von *Synechocystis* sp. PCC 6803

Genomskalige metabolische Netzwerke bieten die Möglichkeit, die physiologischen Eigenschaften des jeweiligen spezifischen Organismus mit *Constraint*-basierten Analysen zu untersuchen [103]. Während für eine Vielzahl von heterotroph wachsenden einzelligen Organismen die Flussbilanzanalyse (*Flux Balance Analysis*) (FBA) schon oft mit großem Erfolg eingesetzt wurde, steht jedoch die Anwendung bei phototroph wachsenden Organismen immer noch am Anfang [145]. Insbesondere ermöglicht das phototrophe Wachstum neue, zusätzliche, mathematische und konzeptionelle Herausforderungen, wie zum Beispiel die Simulation von unterschiedlichen Lichtbedingungen durch verschiedene Stadien eines simulierten Tagesablaufs. Doch bevor solche komplexeren Szenarien berücksichtigt werden können, muss erst eine Referenzlösung für das phototrophe Wachstum unter konstanten Lichtbedingungen gegeben sein.

Für die FBA wird davon ausgegangen, dass die intrazellulären Flüsse so aufgeteilt sind, dass sie eine bestimmte zelluläre Zielfunktion, wie zum Beispiel die Maximierung der Biomasse, gewährleisten. Dabei sind die Rahmenbedingungen, wie auch bestimmte Austauschflüsse, festgesetzt. In der aktuellen Netzwerkversion wurde die BOF aus der Literatur (Nogales et al. [125]) modifiziert (siehe Kapitel 2.1.9). Sie beinhaltet acht unterschiedliche zelluläre Bestandteile: Proteine, DNA, RNA, Zellwandbestandteile, Lipide, lösliche Zellkomponenten, anorganische Ionen und Pigmente. Zusätzlich wurde ein wachstumsabhängiger ATP-Bedarf für den Energiebedarf von Proteinsynthesen, Regulationsprozessen und Wachstum miteinbezogen. Für die Simulationen unter Lichtabwesenheit wurde einzig Glykogen als Speicherstoffverbindung berücksichtigt.

Neben dem Zellwachstum muss die angestrebte FBA-Referenzlösung zusätzliche zelluläre Prozesse widerspiegeln, von denen bekannt ist, dass sie den cyanobakteriellen Stoffwechsel beeinflussen. Dazu gehören unter anderem ein vom Wachstum unabhängiger basaler ATP-Verbrauch für die Aufrechterhaltung zellulärer Prozesse sowie eine verbleibende Restaktivität der Zellatmung, welche für Cyanobakterien auch unter Lichtbedingungen angenommen wird [146, 147]. Des Weiteren sind eine Reihe von anderen Prozessen direkt abhängig vom Lichteinfall, wie die Bildung von reaktiven Sauerstoffradikalen (ROS), Photoinhibition und Schäden durch eine erhöhte Lichteinstrahlung. Um die freiwerdenden Sauerstoffradikale unschädlich zu machen, wird ein erhöhter NADPH-Bedarf benötigt, wohingegen die durch erhöhte Lichtkonzentration verursachten Schäden, vor allem des D1-Proteins des Photosystems II, einen wachstumsunabhängigen Umsatz von zellulären Komponenten zur Folge haben [148].

Erst vor kurzem wurde auch die Existenz einer Mehler-ähnlichen-Reaktion nachgewiesen, bei welcher im Gegensatz zu der Mehler-Reaktion in höheren Pflanzen keine freien Sauerstoffradikale gebildet werden, sondern die Elektronen direkt in Wassermolekülen gebunden werden [149, 150]. Allerdings sind quantitative Daten über diesen

Prozess kaum verfügbar und das Wissen über dessen Wechselbeziehung mit dem Zellwachstum ist nur rudimentär.

Mit dem Ziel das Netzwerkmodell zu parametrisieren, um es mit experimentellen Daten vergleichbarer zu machen, wurde bei den Simulationen unter konstanten Lichtbedingungen eine durchschnittliche Verdopplungszeit von 24 Stunden angenommen. Dies entspricht einer durchschnittlichen Wachstumsrate von $0,029 \text{ h}^{-1}$. Für die Simulation diente Nitrat (NO_3^-) als Stickstoffquelle, welches über ein ABC-Transportersystem aufgenommen wurde. Als einziger wachstumslimitierender Faktor wurde Licht, welches durch den Fluss von absorbierten Photonen repräsentiert wird, angenommen und begrenzt. Andere Nährstoffe, wie zum Beispiel Stickstoff, Phosphor und Schwefel, wurden hingegen nicht limitiert und konnten frei vom System aufgenommen werden. Die Aufnahme von komplexeren Molekülen wie Glucose oder Aminosäuren als Kohlenstoff sowie auch als Stickstoffquelle, wurde bei den Simulationen nicht berücksichtigt. Als Kohlenstoffaufnahme wurde nur Hydrogencarbonat (HCO_3^-) erlaubt, welches über einen aktiven Transport (ABC-Transporter-System) [136] aufgenommen wird. Diese Aufnahme wurde jedoch nicht eingeschränkt. Es wird nur eine Nettoaufnahme von Hydrogencarbonat und kein zyklischer Austauschtransport berücksichtigt [151].

Zusätzlich zu dem wachstumsabhängigen ATP-Bedarf wurde der basale ATP-Verbrauch (*Maintenance*) auf $0,13 \text{ mmol ATP gDW}^{-1} \text{ h}^{-1}$ gesetzt. Um eine basale Aktivität der Zellatmung auch unter Lichtbedingungen einzuberechnen, wurde der terminalen Cytochrom-c-Oxidase also auch der Mehler-ähnlichen-Reaktion, welche NADPH und O_2 zu H_2O und NADP umwandelt, ein Fluss zugewiesen. Es wird angenommen, dass beide Prozesse jeweils ungefähr 10 % des Sauerstoffes, welcher am Photosystem II frei wird, aufnehmen. Hierbei ist zu beachten, dass Schätzungen zur Aktivität dieser beiden Prozesse in der Literatur stark variieren und auch sehr stark von den Wachstumsbedingungen abhängen. Für die Aktivität der terminalen Oxidase werden Werte von Helman et al. (2005) [152] verwendet, der das Ausmaß des Elektronenflusses über Cytochrom-c-Oxidase im Licht beschreibt und auf einen Sauerstoffkonsum von etwa 6 % der Sauerstoffentwicklung schließt. Die Mehler-ähnliche-Reaktion wurde erst vor kurzem von Allahverdiyeva et al. (2011) [150] untersucht, die berichten, dass unter luftähnlichen CO_2 -Bedingungen etwa 20 % der Elektronen, die aus der Wasserspaltung stammen, auf den Sauerstoff übergehen, hauptsächlich auf Grund der Mehler-ähnlichen-Reaktion. Um oxidativen Stress bei erhöhten Lichtkonzentrationen zu modellieren, wurde im Netzwerk eine Reaktion eingebaut, die Superoxide (O_2^-) am Photosystem II und am Photosystem I (durch die von höheren Pflanzen bekannte Mehler-Reaktion) erzeugt. Es wird aber angenommen, dass beide Prozesse nur mit einer geringen Rate ablaufen. Im Netzwerk wird diese so gesetzt, dass ca. 0,5 % der Elektronen im jeweiligen Photosystem direkt auf den Sauerstoff übergehen.

2.2.4. Optimale Flussverteilung für das phototrophe Wachstum

Unter Beachtung der vorher genannten Bedingungen und Einschränkungen wurde mithilfe der COBRA-Toolbox (COBRA) unter MATLAB eine Lösung für die Flussoptimierung unter phototrophen Bedingungen berechnet. Die Flussverteilung liefert die maximale Wachstumsrate zu den vorher definierten externen Bedingungen. Um die erhaltene Lösung zu verifizieren, wurde anschließend mit der FASIMU-Software [107] für die gleiche Simulation unter den gleichen Bedingungen eine Lösung berechnet. Diese Berechnung führte Sabrina Hoffmann durch. Beide Programme lieferten identische Ergebnisse. Zu beachten ist jedoch, dass diese Referenzlösung für einen konstanten Lichteingang nicht eindeutig ist.

Die Flussvariabilitätsanalyse ergab, dass eine deutliche Variabilität in den Reaktionen der Transaldolase (EC 2.2.1.2) (TALDO) (*talB*, slr1793), der Fructose-1,6-Bisphosphat-Aldolase (EC 4.1.2.13) (FBPA) (*cbbA* und *fda*, slr0018 und slr0943) und der Fructose-1,6-Bisphosphatase/ Sedoheptulose-1,7-Bisphosphatase (EC 3.1.3.37) (SBP) (*fbpI* und *glpX*, slr2094 und slr0952) existiert. Diese entsprechen alternativen optimalen Routen innerhalb des Netzwerkes [153]. Eine graphische Übersicht der erhaltenen Flussverteilung ist in Abbildung 2.4 dargestellt. Insgesamt gesehen stimmt diese Flusslösung gut mit den in der Literatur publizierten Ergebnissen überein [124, 125].

Wie erwartet, basiert das autotrophe Wachstum auf der Assimilierung von Kohlenstoffdioxid durch die Ribulose-1,5-Bisphosphat-Carboxylase/-Oxygenase (EC 4.1.1.39) (RuBisCO). RuBisCO wandelt ein Molekül Ribulose-1,5-Bisphosphat (RuBP) und CO₂ in zwei Moleküle Glycerat-3-Phosphat (PG3) um. Um eine weitere Kohlenstofffixierung und damit verbundenes Wachstum zu garantieren, wird PG3 anschließend zum größten Teil genutzt, um RuBP über den Calvin-Benson-Zyklus zu regenerieren. Dabei können für jeweils sechs Moleküle PG3, die vom Zyklus gewonnen werden, ein Molekül PG3 für das Wachstum verwendet werden. Die Energie in Form von ATP und Reduktionsäquivalenten (NADPH), die für diesen Zyklus benötigt wird, werden durch die photosynthetische Lichtreaktion zur Verfügung gestellt. Aus diesem Grund weisen die Reaktionen der Photosynthese sowie des Calvin-Benson-Zyklus die höchsten Flusswerte auf.

Für alle Lösungen gilt, dass innerhalb des Calvin-Benson-Zyklus die Reaktion der Glyceraldehyd-3-Phosphat-Dehydrogenase (EC 1.2.1.12/59) (GAPDH) (*gap2*, slr1342) immer den höchsten, absoluten Fluss besitzt ($230,2 \cdot 10^{-2} \text{ mmol gDW}^{-1} \text{ h}^{-1}$). Diese Reaktion ist eine der Schlüsselreaktionen des Calvin-Benson-Zyklus und essentiell für das phototrophe Wachstum. Die Erzeugung einer entsprechenden Mutante ist demgemäß unter phototrophen Bedingungen nicht möglich (Koksharova et al. [154]). Zu beobachten ist auch, dass über die Reaktion der Ribulosephosphate-3-Epimerase (EC 5.1.3.1) (*cfxE*, slr0807) immer ein höherer Fluss führt als über die Reaktion der Ribose-5-Phosphat-Isomerase (EC 5.3.1.6) (RPI) (*rpiA*, slr0194), mit einer fast doppelt so hohen Flussrate ($81,6$ zu $45,4 \cdot 10^{-2} \text{ mmol gDW}^{-1} \text{ h}^{-1}$). Beide Reaktionen sind im Netzwerk essentiell. Stoffwechselwege, die sich an den Calvin-Benson-Zyklus anschließen und zur Biomassesynthese führen, besitzen eine weit geringere Flussrate.

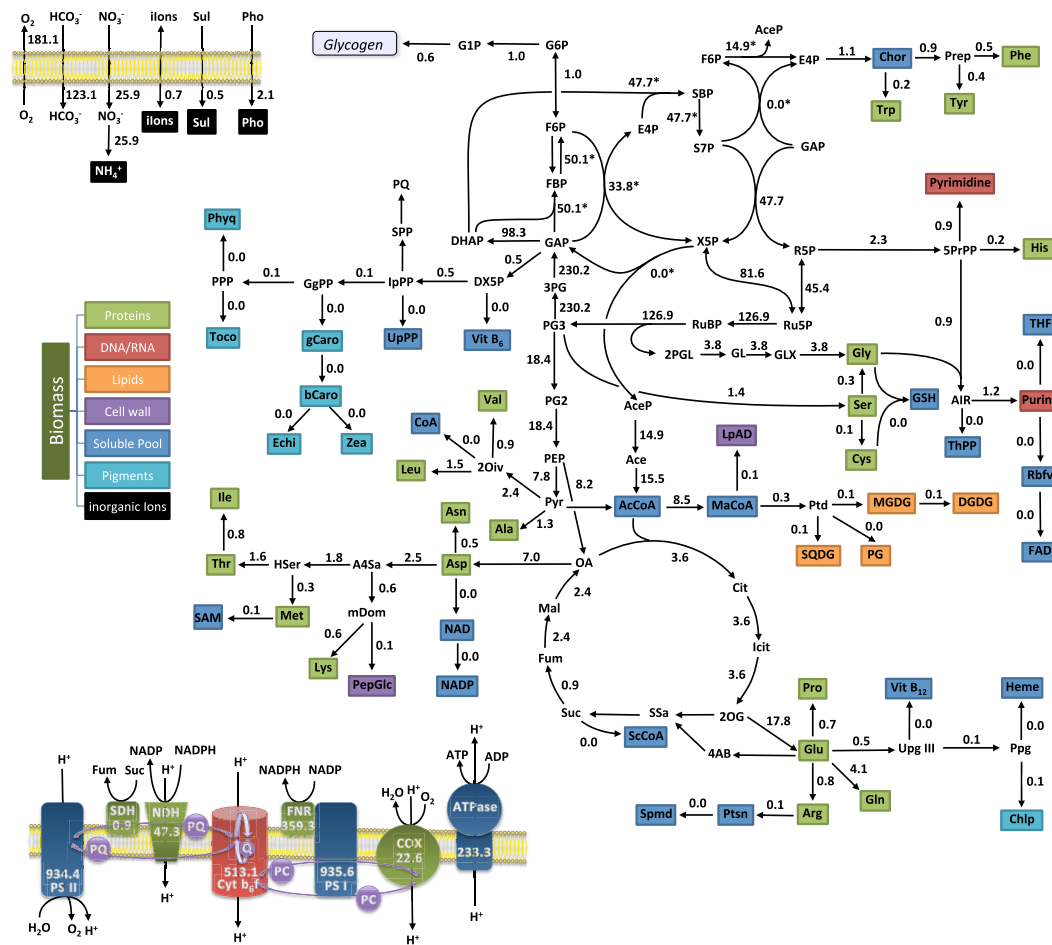


Abbildung 2.4.: Optimale Flussverteilung für *Synechocystis* sp. PCC 6803 unter phototropen Bedingungen. Für die Flussverteilung wurde die maximale Biomasseproduktion optimiert. Die Flüsse sind in $10^{-2} \text{ mmol gDW}^{-1} \text{ h}^{-1}$ gegeben. Variable Flüsse sind durch * markiert. Alle Reaktionen mit einer Flussrate > 0 , aber $< 0,05 \cdot 10^{-2} \text{ mmol gDW}^{-1} \text{ h}^{-1}$ sind durch einen Fluss von 0,0 gekennzeichnet. In der Lösung ist ein hoher Fluss durch den Calvin-Benson-Zyklus zu erkennen und es existiert kein zyklischer Fluss über den TCA-Zyklus. Die Erklärung der Abkürzungen für die verwendeten Metaboliten sind im Anhang in der Tabelle A.3 aufgeführt (Abbildung aus Knoop et al. (2013) [94]).

Während des phototropen Wachstums existiert kein zyklischer Fluss über den TCA-Zyklus, stattdessen dient dieser dazu, Vorläufermetaboliten, wie Aspartat oder 2-Oxoglutarat, für die jeweiligen Synthesewege bereitzustellen. Nur Fumarat, welches als Nebenprodukt unterschiedlicher Stoffwechselwege entsteht, zum Beispiel während der Purinsynthese, wird über den TCA-Zyklus weiter verstoffwechselt.

Bei der Maximierung der Biomassezielfunktion enthält die vorhergesagte Flussverteilung für das phototrophe Wachstum auch die Synthese von Speicherverbindungen, die dann bei der Abwesenheit von Licht wieder abgebaut und genutzt werden können. Um diesen Vorgang etwas zu vereinfachen, wurde bei der Analyse in dieser Arbeit nur Glykogen als einzige Speicherstoffverbindung in die Biomassereaktion aufgenommen.

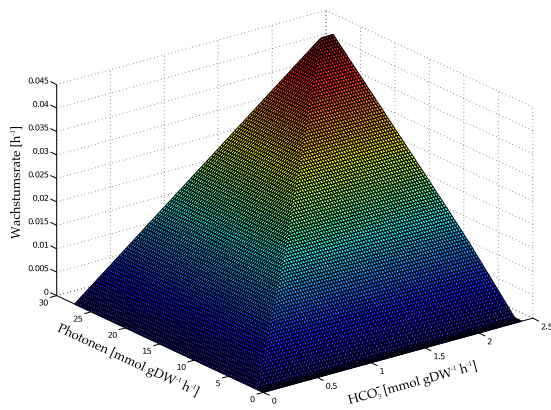


Abbildung 2.5.: Die maximale Wachstumsrate in Abhängigkeit zur Lichtintensität (Photonenanzahl) und HCO_3^- -Verfügbarkeit. Da der festgelegte basale ATP-Verbrauch (*Maintenance*) erfüllt werden muss, ist ein leichter Versatz der Wachstumsrate bei geringen Lichtintensitäten zu erkennen.

In Abbildung 2.5 wird die maximale Wachstumsrate als eine Funktion der Eingangsflüsse Licht (absorbierte Photonen) und HCO_3^- dargestellt. Das maximale Wachstum ist linear abhängig vom absorbierten Licht und der Netto CO_2 -Verfügbarkeit, mit einem Versatz für geringe Lichtintensitäten. Es ist zu erkennen, dass bei einer Variation der Lichtintensität keine qualitativen Unterschiede in der Flussverteilung zu beobachten sind, so lange Licht als der wachstumslimitierende Faktor bleibt.

Der Vergleich von simulierten Modell-
daten mit der Realität ist ein essen-
tieller Schritt in jedem Forschungsbe-
reich. Deshalb wurden auch für diese
Arbeit die Ergebnisse der Flussoptimie-

rung mit experimentellen Daten verglichen. In der Literatur sind unter anderem typische Kohlenstoffaufnahme-
raten für Cyanobakterien zu finden. So haben Van Liere und Walsby [155] eine typische Kohlenstoffaufnahme-
rate von 0 bis 13 $\text{mg Kohlenstoff gDW}^{-1} \text{ h}^{-1}$ gemessen, welche einer Flussrate von ungefähr 0 bis 1,1 $\text{mmol CO}_2 \text{ gDW}^{-1} \text{ h}^{-1}$ entspricht. Bei der Simulation des phototrophen Wachstums, bei der von einer Wachstumsrate von $0,0289 \text{ h}^{-1}$ ausgegangen wird, beträgt der Fluss des Hydrogencarbonat-Transports $1,23 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ und entspricht somit annähernd der gemessenen Rate. Der photosynthetische Quotient, welcher das molare Verhältnis von ausgeschiedenem Sauerstoff und aufgenommenem CO_2 beschreibt, beträgt in der simulierten Flussverteilung ca. 1,47. Des Weiteren wurden 2011 von Young et al. [156] Daten von dynamischen Flussmessungen mithilfe von isoto-
penmarkiertem Kohlenstoff unter photoautotrophem Wachstum von *Synechocystis* publiziert. Die Flussverteilung dieser Messungen stimmt ziemlich genau mit den Ergebnissen der Simulation des phototrophen Wachstums (Abbildung 2.4) überein, gleichwohl es kleine interessante Unterschiede gibt:

Bei der Optimierung der Biomasseproduktion erhält man in der berechneten Lösung einen Fluss durch die Phosphoketolase (EC 4.1.2.22) (PKL), die entweder Xylose-5-Phosphat (X5P) zu Acetylphosphat (ACEP) und Glyceraldehyd-3-Phosphat (GAP) umwandelt oder Fructose-6-Phosphat (F6P) zu Acetylphosphat und Erythrose-4-Phosphat (E4P). Somit ermöglicht die PKL eine Abkürzung vom Calvin-Benson-Zyklus zu Acetyl-CoA und umgeht den CO_2 -Verlust der Reaktion des Pyruvat-Dehydrogenase-Komplexes (PDH). Interessanterweise wurde die Phosphoketolase zuvor schon für den C_5 -Katabolismus in *Saccharomyces cerevisiae* [157] und für die Glutamatproduktion in *Corynebacterium glutamicum* [158] genauer betrachtet und diskutiert. Zudem hat die Anwesenheit der PKL einen signifikanten Einfluss auf den Verzweigungspunkt

von Phosphoglyceratkinase (EC 2.7.2.3) (PGK) und Phosphoglucomutase (EC 5.4.2.1) (PGM), welcher sich an den Calvin-Benson-Zyklus anschließt. Mit einem Fluss durch die PKL erhält man ein Verhältnis von PGK zu PGM von etwa 12,5:1. Ohne das Vorhandensein der PKL beträgt dieses Verhältnis etwa 7,4:1, was auch dem Wert aus früheren Netzwerkversionen [93] und anderen Netzwerkpublikationen [125] entspricht. Das Verhältnis aus den experimentellen Daten von Young et al. [156] entspricht ungefähr $10,0 \pm 1,0$, so dass keine Schlussfolgerung über die Gültigkeit der beiden Lösungen gezogen werden kann. Nichtsdestotrotz liegt der gemessene und simulierte Wert signifikant höher als der Lehrbuchwert, der ein Verhältnis von 5:1 beschreibt. Das liegt daran, dass während der Regeneration des Calvin-Benson-Zyklus Vorstufenmetaboliten für unterschiedliche Biomassekomponenten abgezogen werden, wie zum Beispiel Glucose-6-Phosphat zur Synthese von Glykogen und Nukleotid-Zuckern und Ribose-5-Phosphat zur Synthese von Folat, DNA, RNA und Histidin, die einen höheren Fluss durch die PGK fordern.

Ein weiterer Unterschied zwischen den experimentell bestimmten Daten und der simulierten Flusslösung ist die Rolle des Malatenzyms (ME) während des phototrophen Wachstums. Das ME wandelt Malat zu Pyruvat um und setzt dabei CO_2 frei. In Übereinstimmung mit früheren Studien über heterotrophes Wachstum [140], haben Young et al. [156] einen Fluss durch das ME gemessen, was darauf hindeutet, dass das Enzym bei der intrazellulären Konzentration von CO_2 beteiligt sein könnte, analog zu seiner Rolle in C_4 -Pflanzen. Da jedoch ein solcher Zyklus Energie in Form von ATP konsumiert, wird bei der Flussoptimierung durch die FBA kein solcher Fluss in einer Lösung vorhanden sein, solange nicht noch weitere Einschränkungen oder Nebenbedingungen festgelegt werden.

Die Flussmessungen von Young et al. (2011) [156] zeigen einen geringen Fluss durch den oxidativen Pentosephosphatweg (OPP), welcher unter der Freisetzung von CO_2 zur Erzeugung von NADPH beiträgt. Dieser ist unter phototrophen Bedingungen nicht zu erwarten, da unter Anwesenheit von Licht ein Großteil des NADPHs durch die Elektronentransportkette (ETK) erzeugt wird, wodurch ein Fluss durch den OPP zu einem geringen aber signifikanten Verlust von CO_2 führt. Young et al. [156] argumentieren, dass dieser geringe Fluss aus einer unvollständigen Unterdrückung der Enzymaktivität des OPP während des Tages resultiert. Jedoch kann eine solche suboptimale Flusslösung nicht ohne weitere Beschränkungen beziehungsweise ein explizites Fordern durch eine FBA simuliert werden.

Tatsächlich wurde schon gezeigt, dass sich der Zellstoffwechsel in unterschiedlichen Standby-Modi befinden kann, um schnell und direkt auf sich ändernde Umwelteinflüsse zu reagieren, was auf Kosten des optimalen Wachstums geht [110, 159]. Dieser Kompromiss zwischen Flexibilität und Effizienz erfordert die Investition von zusätzlichen Ressourcen und spielt sehr wahrscheinlich eine wichtige Rolle bei der Regulation des Stoffwechsels von phototrophen Organismen beim Wachstum unter natürlichen Tag-Nacht-Bedingungen.

2.2.5. Die Oxygenase-Reaktion von RuBisCO und die Photorespiration

Das bifunktionale Enzym Ribulose-1,5-Bisphosphat-Carboxylase/-Oxygenase (EC 4.1.1.39) besitzt nicht nur eine Affinität für CO₂, sondern ebenfalls noch für molekularen Sauerstoff (O₂), welches es als alternatives Substrat nutzen kann. Bei der Carboxylierung von RuBisCO entstehen zwei Moleküle Glycerat-3-Phosphat (PG3), wohingegen bei der konkurrierenden Oxygenierung ein Molekül PG3 und ein Molekül 2-Phosphoglycolat (2PGL) entstehen (Photorespiration), wovon letzteres nur aus zwei Kohlenstoffatomen zusammengesetzt ist und sie somit ineffektiver zur Carboxylierung macht.

Betrachtet man den heutigen Zustand der Erdatmosphäre, mit ca. 21 % Sauerstoff und ca. 0,04 % CO₂, wird deutlich, dass die Produktivität von photosynthetischen Organismen signifikant durch die Biofunktionalität von RuBisCO beeinträchtigt wird. Diese Tatsache wird meist als ein evolutionäres Relikt aus der Zeit gesehen, in der sich die Photosynthese beziehungsweise RuBisCO entwickelte. Im Vergleich zu heute herrschte zu dieser Zeit eine viel höhere CO₂-Konzentration in der Atmosphäre und die O₂-Konzentration war minimal. Infolgedessen wird auch die Photorespiration als ein scheinbar „verschwenderischer“ Prozess gesehen, der Kohlenstoff aus dem Calvin-Benson-Zyklus entnimmt. Dementgegen existieren aber auch Hypothesen, die darauf hindeuten, dass die Photorespiration eine wesentliche Funktion bei der Energiedissipation bei hohen Lichtkonzentrationen [160] oder anderen Prozessen spielt [161].

Die Rolle und Bedeutung der Photorespiration in Cyanobakterien ist bis heute nur teilweise geklärt. Es ist bekannt, dass Cyanobakterien CO₂-Konzentrationsmechanismen zur lokalen Erhöhung der CO₂-Konzentration entwickelt haben, wodurch sie in der Lage sind, die Oxygenase-Aktivität von RuBisCO wesentlich zu unterdrücken und dadurch die photosynthetische Leistung zu steigern [162]. Allerdings ermöglichen die CO₂-Konzentrationsmechanismen nicht die völlige Unterdrückung der Oxygenierung und Eisenhut et al. (2008) [119] haben gezeigt, dass ein aktiver photorespiratorischer 2-Phosphoglycolat-Stoffwechsel für das cyanobakterielle Wachstum in der heutigen sauerstoffhaltigen Atmosphärenzusammensetzung essentiell ist. Insbesondere wird eine schnelle Umwandlung von 2PGL zu Glyoxylat bei einer nicht unwesentlichen Photorespirationsrate benötigt, um einen toxischen Effekt und eine Hemmung des Calvin-Benson-Zyklus zu verhindern.

Als überraschendes Ergebnis zeigten sich in der ersten Version der Netzwerkrekonstruktion inklusive deren Optimierungslösungen für den *Synechocystis*-Metabolismus, dass bei der Maximierung von Biomasse unter phototrophen Bedingungen ein geringer Fluss über die Oxygenierungsreaktion von RuBisCO vorhergesagt wird. Dieses Ergebnis lässt sich auch in experimentell bestimmten Daten wiederfinden [93]. Der Grund für diesen scheinbar nicht optimalen Stoffwechselweg war das Nichtvorhandensein eines stöchiometrisch effektiveren Weges für die Synthese der Aminosäuren Serin, Glycin und Cystein. Insbesondere konnten für die Enzyme Phosphoserin-Transaminase (EC 2.6.1.52) und Phosphoserin-Phosphatase (EC 3.1.3.3) zunächst keine Homologe im Genom von *Synechocystis* gefunden werden.

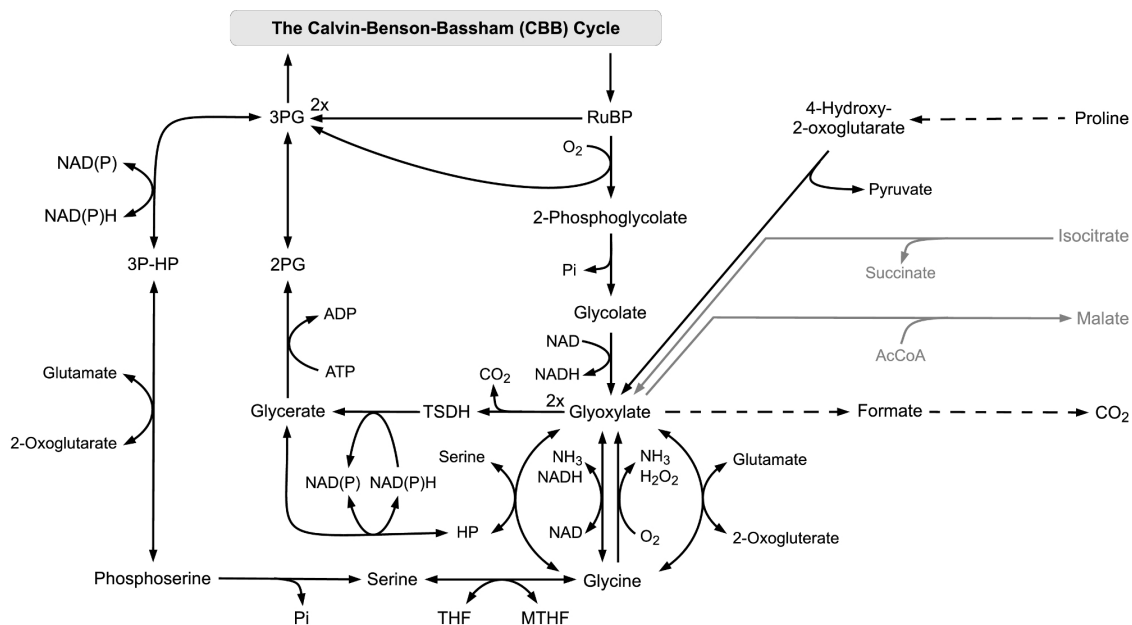


Abbildung 2.6.: Die Stoffwechselwege der Photorespiration und des Glyoxylat-Stoffwechsels in *Synechocystis*. Zusätzlich ist die Serin-Synthese über den Phosphoserin-Weg gezeigt. Die Erklärung der Abkürzungen für die verwendeten Metaboliten sind im Anhang in der Tabelle A.3 aufgeführt (Abbildung aus Knoop et al. (2013) [94]).

Erst vor kurzem wurde entdeckt, dass ein Homolog für eine Glucomutase ein vielversprechender Kandidat für die Aktivität der Phosphoserin-Phosphatase ist [131]. Desgleichen konnte die Aktivität einer Phosphoserin-Transaminase in *Synechocystis* bestätigt werden (persönliche Mitteilung Wolfgang Lockau und Martin Hagemann). Genauere Untersuchungen hierfür stehen aber noch aus.

Ohne den Phosphoserin-Weg wird Glycin aus Glyoxylat synthetisiert, welches als Nebenprodukt aus der Photorespiration entsteht. Aus diesem Grund wurden für diese Arbeit drei Szenarien im Zusammenhang mit den Synthesemöglichkeiten der Aminosäuren Glycin, Serin und Cystein genauer untersucht.

Beim ersten Szenario wird davon ausgegangen, dass die derzeitige Annotation vollständig ist und dass kein kompletter Phosphoserin-Weg in *Synechocystis* existiert. Für diesen Fall werden Glycin, Serin und Cystein aus Glyoxylat synthetisiert. Glyoxylat ist selbst wiederum ein Produkt aus Glycolat und somit auch aus 2PGL, das Produkt der Photorespiration. Dies hat somit auch zur Folge, dass es bei der Flussoptimierung einen geringen Fluss durch die Photorespiration gibt (siehe Abbildung 2.6 für eine Übersicht der Stoffwechselwege). Es ergibt sich eine Photorespirationsrate von ungefähr 3,5 %, welche gut mit aktuellen Schätzungen der *in vivo* Rate übereinstimmt [163]. Jedoch ist zu beachten, dass Young et al. [156] eine deutlich niedrigere Rate der Photorespiration unter ähnlichen experimentellen Bedingungen beobachteten. Sieht man sich unter diesen Gesichtspunkten eine Simulation des Stoffwechsels unter Abwesenheit von Licht genauer an, ist ein Fluss durch die Photorespiration bei der optimalen Flussverteilung nicht mehr gegeben. Hingegen wird der verbleibende

Bedarf an Glycin, Serin und Cystein durch den Abbau von Prolin gewährleistet, welches effektiver ist als die Gewinnung von Glyoxylat durch die Photorespiration unter nicht phototrophen Bedingungen.

Als zweites Szenario wird angenommen, dass noch zum Teil nicht identifizierte Gene für eine Phosphoserin-Transaminase und -Phosphatase kodieren und damit der Phosphoserin-Weg komplett wäre. Werden nun die zugehörigen enzymatischen Schritte beziehungsweise Reaktionen in das Modell eingebracht, erhält man keinen Fluss mehr durch die Reaktionen der Photorespiration unter phototrophen Bedingungen und die Oxygenierungsreaktion von RuBisCO ist nicht mehr Teil der optimalen Lösung. Stattdessen werden Serin und anschließend Glycin und Cystein nun durch den neu eingeführten Phosphoserin-Weg synthetisiert. Experimentell lässt sich dieses Szenario durch Ergebnisse von Young et al. [156] und Huege et al. [163] festigen, da beide Studien eine wesentlich höhere ^{13}C -Anreicherung bei Serin gegenüber Glycin messen. Hieraus lässt sich schließen, dass zuerst Serin synthetisiert wird und danach erst Glycin.

Für das dritte und letzte Szenario wird angenommen, dass die Photorespiration ein unvermeidbarer Prozess unter den gegenwärtigen atmosphärischen Bedingungen ist. Es wurde die optimale Flussverteilung unter Miteinbeziehung des Phosphoserin-Weges in das Modell und gleichzeitiger Festlegung eines Flusses durch die Oxygenierungsreaktion von RuBisCO berechnet. Zu diesem Zweck wurde die Carboxylierungs- mit der Oxygenierungsreaktion gekoppelt. Das Verhältnis dieser beiden Reaktionen wurde auf 3 % Oxygenierung zu 97 % Carboxylierung festgelegt, so dass immer ein minimaler Fluss über 2PGL existieren muss. Für dieses Szenario zeigt die Flussvariabilitätsanalyse, dass mehrere gleichwertige, optimale Lösungen existieren. Zum Beispiel gibt es für das Glyoxylat, welches als Zwischenprodukt der Photorespiration entsteht, mehrere Abbaumöglichkeiten. Es kann entweder zur Synthese von Glycin für die Biomasse verwendet werden oder stöchiometrisch äquivalent (in Bezug auf die Biomasse) über Glycerat verstoffwechselt werden. Im letztgenannten Fall wird Serin über den Phosphoserin-Weg synthetisiert (siehe Abbildung 2.6).

Um schlussendlich eine mögliche Aussage über die Existenz des Phosphoserin-Weges in *Synechocystis* zu treffen, ist zu bedenken, dass die Photorespiration als unvermeidlicher Prozess unter den aktuellen atmosphärischen Bedingungen gilt und es von Vorteil ist, Produkte, die dabei zwangsweise entstehen, stöchiometrisch so effizient wie möglich zu nutzen. Zudem wurde schon 1997 vorgeschlagen, dass dieser Weg in Cyanobakterien vorhanden sei [164] und zusätzlich weisen die ^{13}C -Markierungsexperimente von Young et al. [156] auf eine Existenz des Phosphoserin-Weges in *Synechocystis* hin.

Desgleichen ist ein Enzym des Phosphoserin-Weges, die 3-Phosphoglycerat-Dehydrogenase (EC 1.1.1.95), in mehreren metabolischen Datenbanken annotiert. Zwar haben Eisenhut et al. [119] dieses Gen (sll1908) als Hydroxypyruvat-Reduktase annotiert, allerdings deuten jüngste Experimente daraufhin, dass das Gen tatsächlich für eine 3-Phosphoglycerat-Dehydrogenase kodiert. Ferner konnte ein alternatives Gen für die Hydroxypyruvat-Reduktase identifiziert werden (persönliche Mitteilung Martin Hagemann). Schlussendlich zeigen erste experimentelle Messungen, dass alle

Enzyme des Phosphoserin-Weges mit der entsprechenden Funktion in *Synechocystis* exprimiert werden und die gesuchten Reaktionen katalysieren können (siehe Kapitel 2.1.6). Aus diesen vorher genannten Gründen wird dieses letzte Szenario, in welchem der Phosphoserin-Weg aktiv ist und eine nicht vermeidbare Photorespiration existiert, präferiert und in die Rekonstruktion und das mathematische Modell des Stoffwechsels von *Synechocystis* integriert, da es wohl am ehesten der *in vivo* Situation im Organismus entspricht.

2.2.6. Optimale Flussverteilung bei Abwesenheit von Licht

Neben dem phototrophen Wachstum bei Tageslicht müssen Cyanobakterien auch in der Lage sein, längere Perioden der Dunkelheit, zum Beispiel bei Nacht, zu überleben. Dabei greifen sie auf endogene Speicherverbindungen wie Glykogen zurück und auch zu einem geringen Maße auf Cyanophycin und Poly- β -Hydroxybutyrat, die sich in den Zeitabschnitten angesammelt haben, als Licht zur Verfügung stand. Leider sind experimentell erworbene Daten von Flussverteilungen bei längerer Dunkelheit kaum in der Literatur verfügbar.

Um sich dennoch einem solchen Szenario anzunähern und eine Flussverteilung zu beschreiben, wurde vermutet, dass anders als bei einigen diazotrophen Cyanobakterien, welche einen sehr aktiven Dunkelstoffwechsel betreiben, bei *Synechocystis* nur eine sehr geringe Aktivität des Stoffwechsels besteht und diese sich hauptsächlich auf die Erzeugung von Energie für den zellulären basalen ATP-Verbrauch (*Maintenance*) beschränkt. Es wird daher davon ausgegangen, dass das zelluläre Wachstum bei Dunkelheit minimal ist, was auch schon durch zahlreiche experimentelle Beobachtungen bestätigt wurde.

Aus diesem Grund wurde bei der Flussoptimierung bei Abwesenheit von Licht angenommen, dass die maximale Glykogenabbaurate geringfügig über dem Wert lag, der benötigt wird, um die Vorgaben für die Zellatmung und den basalen ATP-Verbrauch zu erfüllen und dementsprechend festgelegt. Zusätzlich wurde angenommen, dass der restliche, minimale Anteil des Glykogens, welches durch den Abbau zur Verfügung stand, für das Zellwachstum oder äquivalent als gesamt zellulärer Metabolitumsatz, der ebenfalls durch die Biomassefunktion angenähert wird, genutzt wird [93]. Kurz gesagt, die Optimierung für eine Dunkelperiode versucht die BOF bei einer Beschränkung des Glykogenabbaus zu maximieren, der nur geringfügig oberhalb der Anforderung für den basalen ATP-Verbrauch (*Maintenance*) liegt. Glykogen wird als das Hauptsubstrat für die Zellatmung während der Dunkelheit angesehen. Es ist aber darauf hinzuweisen, dass *Synechocystis* sehr wahrscheinlich auch in der Lage ist, andere Speicherverbindungen zu nutzen. Kürzlich wurde gezeigt, dass Mutanten mit einer unterbrochenen Glykogensynthese eine reduzierte Lebensfähigkeit unter Tag-Nacht-Bedingungen aufweisen. Diese fällt jedoch nicht so stark aus, dass man daraus schließen könnte, dass Glykogen die einzige Speicherstoffverbindung ist [165]. Bei den folgenden Analysen wurde der Einfluss von alternativen Speicherverbindungen nicht berücksichtigt.

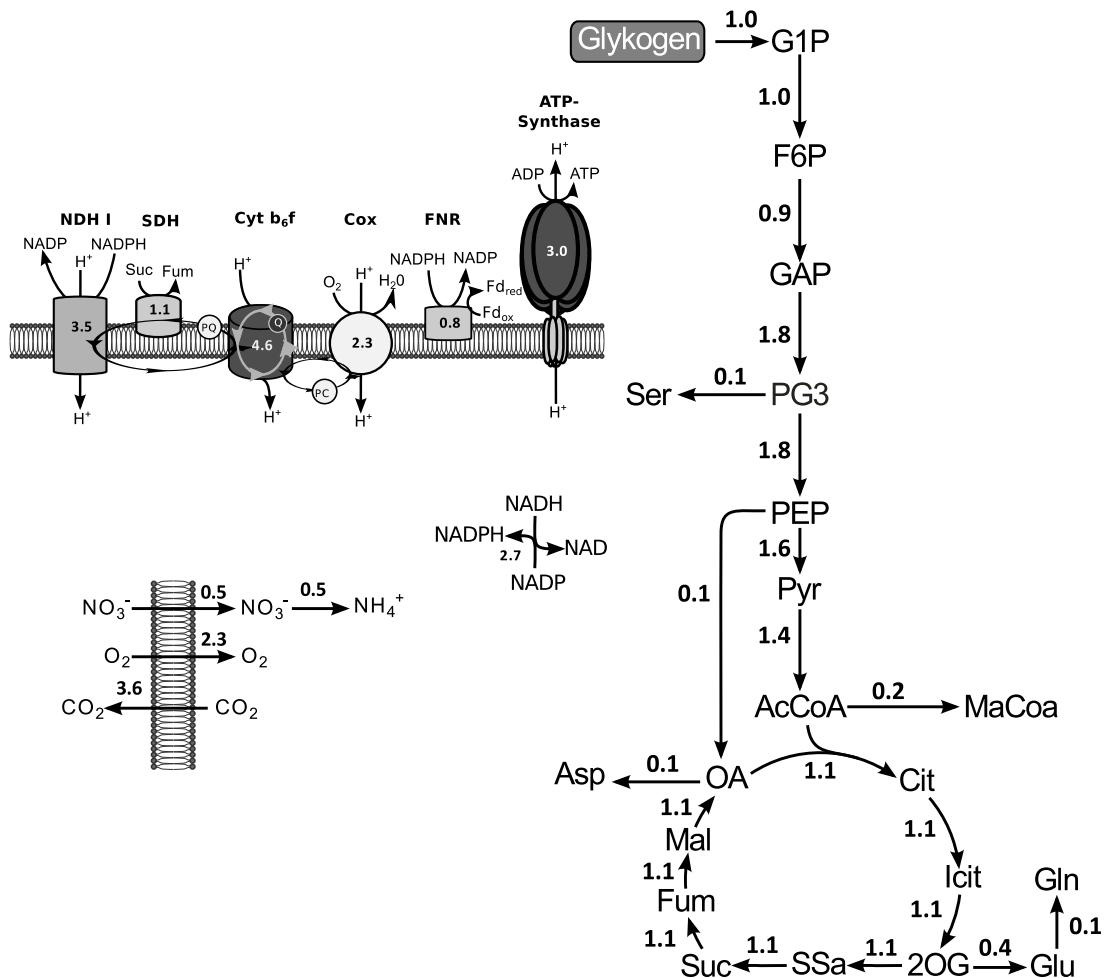


Abbildung 2.7.: Ausschnitt des Zentralstoffwechsels der optimalen Flussverteilung für *Synechocystis* sp. PCC 6803 bei Abwesenheit von Licht. Es ist ein zyklischer Fluss über den TCA-Zyklus zu erkennen. Die Flüsse sind in 10^{-2} mmol gDW⁻¹ h⁻¹ gegeben. Die Erklärung der Abkürzungen für die verwendeten Metaboliten sind im Anhang in der Tabelle A.3 aufgeführt.

Die resultierende optimale Flussverteilung zeigt eine große Variabilität bezogen auf die Annahmen über die Spezifität beziehungsweise Reaktionsrichtung bestimmter Enzyme. Hierbei ist vor allem die Reaktion der annotierten Transhydrogenase (slr1239 und slr1434; EC 1.6.1.2) zu nennen. Wird davon ausgegangen, dass dieses Enzym im Stoffwechsel aktiv ist und einen reversiblen Fluss katalysieren kann, so wird das Redoxpotenzial (NADH) hauptsächlich über den TCA-Zyklus generiert und hinterher zu NADPH umgewandelt. Das gewonnene NADPH wird anschließend zum größten Teil über den NADPH-Dehydrogenasekomplex (NDH I) für die Zellatmung verbraucht, für den gezeigt wurde, dass dieser spezifisch für NADPH ist [146].

Wiederum eine ganz andere Flussverteilung ergibt sich, wenn man davon ausgeht, dass entweder die Transhydrogenase nicht vorhanden ist oder nur unidirektional die

Reaktion von NADPH zu NADH katalysiert. Für diesen Fall wird in der errechneten optimalen Flussverteilung das Redoxpotenzial für die Zellatmung (NADPH) überwiegend über den oxidativen Pentosephosphatweg (OPP) und nicht zyklisch über den TCA-Zyklus generiert. Ein Fluss über den OPP würde gut mit experimentell gewonnenen Daten übereinstimmen, die beim heterotrophen Wachstum voraussagen, dass ein großer Anteil der aufgenommenen Glucose über den OPP oxidiert wird [140]. Erlaubt man nun hingegen, dass die NADH I die beiden Redoxpotenziale NADH und NADPH gleichermaßen als Substrat nutzen kann, erhält man in der errechneten Flussverteilung erneut einen zyklischen Fluss durch den TCA-Zyklus. Die Verwendung des zyklischen TCA-Zyklus wird ebenfalls durch die Beobachtungen und Messungen gestützt, wonach die Succinat-Dehydrogenase (EC 1.3.99.1) (SDH) (sll0823, sll1625, slr1233) den Hauptumsatz an Elektronen in die Elektronentransportkette beziehungsweise den Plastochinon-Pool liefert [146]. Schlussendlich ist es auf der Grundlage der gegenwärtig verfügbaren Informationen leider nicht möglich, eine endgültige Aussage über die tatsächliche Flussverteilung bei Abwesenheit von Licht zu treffen. Da auch keine zusätzlichen Informationen zur Transhydrogenase-Aktivität bekannt sind, wurde in dieser Arbeit das Szenario mit dem zyklischen TCA-Zyklus bevorzugt. Die simulierte Flussverteilung unter Abwesenheit von Licht ist in Abbildung 2.7 gezeigt.

2.2.7. Übergang von der Nutzung von Speicherstoffen hin zum phototropen Wachstum

Durch die Simulation eines Überganges von Tag- zu Nacht-Bedingungen, ermöglicht das *in silico* Modell die Beobachtung, wie sich der Stoffwechsel vom Glykogen-verwendenden Nachtstoffwechsel zum phototropen Wachstum reorganisiert und wie sich dabei bestimmte Reaktionen im Netzwerk verhalten. Dieser metabolische Übergang ist in Abbildung 2.8 für einige ausgewählte Reaktionen gezeigt. Hierbei werden die Flüsse farbkodiert auf einer zweidimensionalen Ebene als eine Funktion der absorbierten Photonen (Abszisse) und der Glykogen-Nutzung (Ordinate) gezeigt. Innerhalb jedes Einzelgraphen entspricht der Übergang vom Nacht-Stoffwechsel (Glykogen-Mobilisierung) zum phototropen Wachstum einem Weg von der oberen linken zur unteren rechten Ecke. Einher eines solchen Verlaufs sind mehrere charakteristische Übergänge in der Reorganisation der metabolischen Flüsse zu beobachten. Bei der Verwendung der Speicherstoffe, was einem Punkt auf der Ordinate in Abbildung 2.8 entspricht, wird angenommen, dass Glykogen als einzige Ressource für Energie und Kohlenstoff fungiert. Für diesen Fall ist die Flussverteilung dadurch gekennzeichnet, dass eine Sauerstoffaufnahme vorliegt und ein Fluss durch die Reaktion der RuBisCO nicht stattfindet. Stattdessen ist ein zyklischer Fluss durch den TCA-Zyklus (Abbildung 2.8 G) über die Succinat-Dehydrogenase (EC 1.3.99.1) (SDH) (slr037) zu beobachten.

Mit zunehmender Lichtintensität und gleichzeitiger Abnahme der Glykogen-Nutzung, nimmt dieser Fluss durch den TCA-Zyklus jedoch ab und stattdessen gewinnt der OPP an Aktivität (Abbildung 2.8 H). Bei steigenden Lichtverhältnissen beginnen die photosynthetischen Lichtreaktionen die Erzeugung der Energie (in Form von ATP) und Redoxäquivalente (NADPH), die für den basalen ATP-Verbrauch (*Maintenance*)

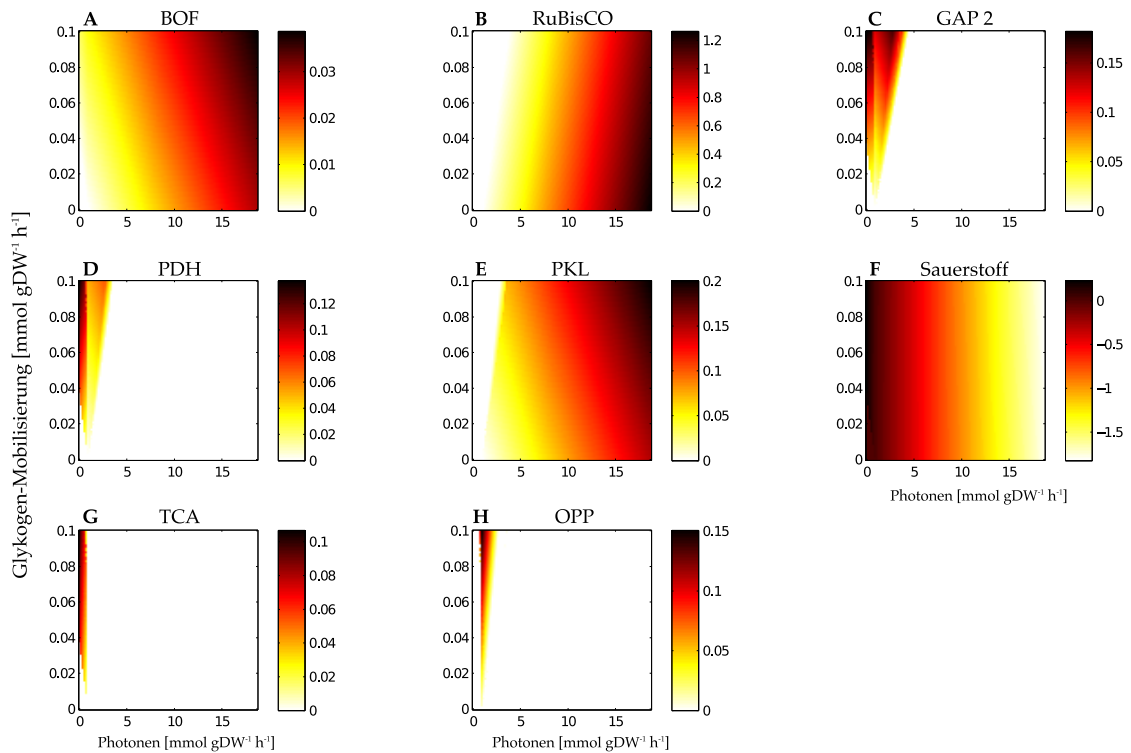


Abbildung 2.8.: Übergang von der Nutzung von Speicherstoffen hin zum phototrophen Wachstum. Gezeigt ist der Fluss durch die entsprechenden Reaktionen (farbkodiert) als Funktion der absorbierten Photonen (Abszisse) und der Glykogen-Mobilisierung (Ordinate). Alle Flüsse sind in $\text{mmol gDW}^{-1} \text{h}^{-1}$ gegeben, bis auf die Biomassezielfunktion (A). Die Netto-Sauerstoffaufnahme (F) nimmt positive Werte entsprechend der Zellatmung an. A) BOF: Biomassezielfunktion B) RuBisCO: RuBisCO-Reaktion C) GAP2: Glyceraldehyd-3-Phosphat-Dehydrogenase 2 D) PDH: Pyruvat-Dehydrogenase-Komplex E) PKL: Phosphoketolase F) Sauerstoffaufnahme G) TCA: zyklischer Fluss über den TCA-Zyklus H) OPP: oxidativer Pentosephosphatweg.

und für die CO_2 -Fixierung benötigt werden. Somit wird ab einem bestimmten Schwellenwert ein Fluss durch den Calvin-Benson-Zyklus eingeleitet, was zu einer zunehmenden Rate der Carboxylierungs- und Oxygenierungsreaktion von RuBisCO führt und gleichzeitiger Aktivität der Phosphoketolase (Abbildung 2.8 B und E). Diese ermöglicht es den CO_2 -freisetzenden Schritt des Pyruvat-Dehydrogenase-Komplexes zu umgehen (Abbildung 2.8 D). Zusätzlich steigt ab diesem Zeitpunkt die Sauerstoffabgabe aufgrund der photosynthetischen Lichtreaktionen (Abbildung 2.8 F).

2.2.8. Der cyanobakterielle Glyoxylat-Shunt

Die meisten der in der Literatur veröffentlichten Netzwerkrekonstruktionen des metabolischen Netzwerkes von *Synechocystis* beinhalten die Stoffwechselreaktionen der Isocitratlyase (EC 4.1.3.1) (ICL) und der Malat-Synthase (EC 2.3.3.9) (MS). Hieraus erhält man einen funktionierenden, bakteriellen Glyoxylatzyklus [120, 122–124].

Die Isocitratlyase, das erste Enzym dieses Zyklus, spaltet Isocitrat zu Succinat und Glyoxylat. Die Malatsynthase verbindet anschließend Glyoxylat und Acetyl-CoA zu Malat und Coenzym A. Jedoch sind die entsprechenden Gene für beide Enzyme nicht im Genom annotiert. Das Einfügen dieser beiden Reaktionen in das entsprechende Netzwerk wurde zum Teil damit begründet, dass es Anzeichen gab, dass die jeweilige enzymatische Aktivität experimentell nachgewiesen werden konnte [140, 166, 167]. Allerdings sind diese experimentellen Ergebnisse nicht schlüssig [93]. Überdies wurden kürzlich erst Gene für einen funktionierenden Glyoxylatzzyklus im Genom von einigen *Cyanothece*-Stämmen identifiziert [168]. Diese besitzen jedoch keine Homologie zu Genen im Genom von *Synechocystis*.

Um Gewissheit in dieser Sache zu bekommen, wurden in dieser Arbeit die funktionalen Auswirkungen eines Glyoxylatzzyklus auf die optimale Flussverteilung beim phototrophen Wachstum durch mehrere Szenarien getestet. Zusätzlich wurden in Kooperation mit der Arbeitsgruppe ‚Biochemie der Pflanzen‘, die einzelnen enzymatischen Schritte eines Glyoxylatzzyklus bei *Synechocystis* experimentell untersucht und validiert. Die zugehörigen Experimente wurden von Marianne Gründel durchgeführt.

Wenn eine Isocitratlyase in die Stoffwechselrekonstruktion von *Synechocystis* eingebracht wird, ist deren enzymatische Reaktion ein Teil der berechneten optimierten Flussverteilung. In diesem Fall dient die Isocitratlyase zur Synthese von Glyoxylat, welches als Vorstufe für die Aminosäuren Glycin, Serin und Cystein fungiert. Demzufolge sind die Reaktionen der Photorespiration bei dieser Flusslösung unter diesen Bedingungen dafür nicht notwendig. Wenn jedoch angenommen wird, dass der Phosphoserin-Weg existiert und funktioniert, gibt es weder einen Fluss durch den kompletten Glyoxylatzzyklus noch durch die Isocitratlyase. Für die Simulation des Zellwachstums unter der Abwesenheit von Licht erhält man ebenfalls einen Fluss in der Optimierung durch den kompletten Glyoxylatzzyklus. Es lässt sich noch anmerken, dass in einigen Netzwerkrekonstruktionen [120] der Glyoxylatzzyklus eingeführt wurde, um den sonst unvollständigen cyanobakteriellen TCA-Zyklus zu schließen.

Um experimentell die Möglichkeit einer enzymatischen Aktivität der Isocitratlyase in *Synechocystis* nachzuweisen, wurde eine verbesserte Methode zur Isocitratlyase-Aktivitätsbestimmung im zellfreien Extrakt angewandt. Zu diesem Zweck wurde zunächst die Methode von Dixon und Kornberg [169] adaptiert. Diese beschreibt ein Standardverfahren für den Nachweis und die Bestimmung von Glyoxylat in löslichen Zellextrakten. Hierbei reagiert Glyoxylat mit Phenylhydrazin zu Phenylhydrazon, welches durch Absorption bei 324 nm gemessen werden kann. Zu beachten ist jedoch, dass auch andere Metaboliten mit reaktiven Keto- oder Aldehydgruppen zu der gleichen Reaktion führen können. Deswegen sei besonders 2-Oxoglutarat erwähnt, welches durch die Decarboxylierung von Isocitrat durch die Isocitrat-Dehydrogenase in Anwesenheit von NADP entsteht.

Aus diesem Grund, wurden die Messungen in Rohextrakten von *Synechocystis*-Zellen durchgeführt, die über eine PD-10 Gelfiltrationssäule aufgereinigt wurden, um kleine reaktive Metaboliten und NADP zu entfernen. Die Ergebnisse ließen erkennen, dass keine signifikante Isocitratlyase-Aktivität in gefilterten Rohextrakten von

Synechocystis nachweisbar war (Abbildung 2.9 A; Spur B). Ohne den Filtrationsschritt konnte eine kleine Erhöhung der Absorption bei 324 nm auch in Abwesenheit von Isocitrat gemessen werden. Diese beruht wahrscheinlich auf der Absorption von kleinen reaktiven Metaboliten im Rohextrakt. Nach Zugabe von Isocitrat konnte eine viel höhere Zunahme der Absorption (324 nm) gemessen werden (Abbildung 2.9 A; Spur A), welche ähnlich zu den Messungen aus der Literatur sind [166, 167].

Diese Ergebnisse zeigen, dass die gemessene Zunahme der Absorption (324 nm) in Rohextrakten, die nicht durch Gelfiltration aufgereinigt wurden, vor allem aus der Bil-

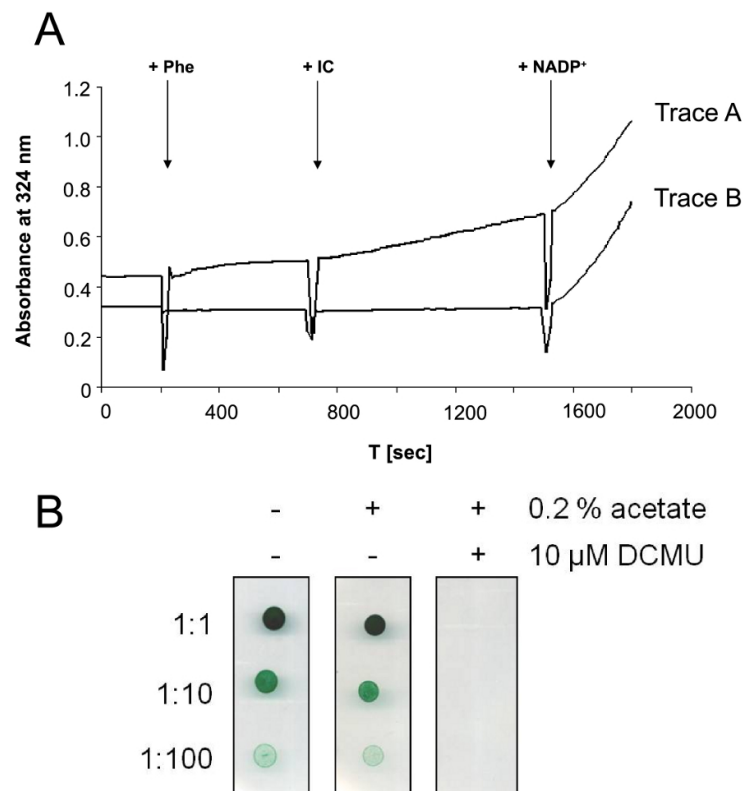


Abbildung 2.9.: Experimentelle Überprüfung des Glyoxylat-Shunts.

A) Enzymatischer Test um die Isocitratlyase-Aktivität in zellfreien Extrakten von *Synechocystis* sp. PCC 6803 zu bestimmen. Messung der Phenylhydrazon-Bildung bei A324 nm. Spur A: Messung im Rohextrakt. Spur B: Messung im aufgereinigten Rohextrakt. Die Pfeile markieren die schrittweise Zugabe von Phenylhydrazin (Phe, 5 mM), Isocitrat (IC, 1 mM) und NADP (0,2 mM). Der Anstieg der A324 nm in der Spur A nach der Zugabe von Phenylhydrazin lässt sich durch reaktive Metaboliten erklären. Der weitere Anstieg nach der Zugabe von IC erklärt sich durch internes NADP im Rohextrakt, welches durch die Isocitrat-Dehydrogenase mit dem IC zu 2-Oxoglutarat reagiert. Beide Spuren zeigen den gleichen starken Anstieg nach der externen Zugabe von NADP, der sich auf die Aktivität der Isocitrat-Dehydrogenase zurückführen lässt.

B) Spot-Assay zur Überprüfung der Acetat-Aufnahme und der Nutzung von Acetat als alleinige Kohlenstoffquelle. Die photosynthetische Aktivität wurde mittels des Photosystem II-Inhibitors DCMU unterdrückt, um die Fähigkeit des photoheterotrophen Wachstums zu bestimmen. Die entsprechenden Kontrollen wurde ohne DCMU und mit oder ohne externes Acetat durchgeführt. Gezeigt sind verschiedene Verdünnungsstufen (1:1, 1:10 und 1:100) von *Synechocystis*-Kulturen auf BG11-Agar (Experimentelle Ergebnisse von Marianne Gründel und Yvonne Zilliges; Abbildung aus Knoop et al. (2013) [94]).

dung von 2-Oxoglutarat resultiert. Verifiziert wurde dies außerdem durch einen enzymatischen Test für 2-Oxoglutarat (nicht dargestellt). Daraus lässt sich schließen, dass die Isocitratlyase-Aktivität, die in der Literatur beschrieben wird, sehr wahrscheinlich das Ergebnis der Isocitrat-Dehydrogenase-Reaktion ist.

Weitere Hinweise auf das Nichtvorhandensein des Glyoxylatzyklus gibt die Tatsache, dass *Synechocystis* nicht die Fähigkeit besitzt, Acetat als alleinige Kohlenstoffquelle in Gegenwart des Photosystem II-Inhibitors 3-(3,4-Dichlorphenyl)-1,1-Dimethylharnstoff (DCMU) (photoheterotrophes Wachstum) zu nutzen. Dieser Test wurde von Yvonne Zilliges durchgeführt. Simulationen mit dem Netzwerkmodell zeigen hier, dass beim Vorhandensein des Glyoxylatzyklus, ein solches photoheterotrophes Wachstum möglich wäre. Die Versuchsergebnisse von Marianne Gründel zum Nachweis einer Isocitratlyase-Aktivität und der Spot-Assay zur Überprüfung der Acetat-Aufnahme von Yvonne Zilliges sind in der Abbildung 2.9 B gezeigt.

2.2.9. Der TCA-Zyklus der Cyanobakterien

Als einer der wichtigsten Wege im zentralen Stoffwechsel hat der TCA-Zyklus eine zweifache Rolle. Zum einen wird er für die Oxidation von Substraten für die Atmungskette zur ATP-Generierung gebraucht und zum anderen für die Bereitstellung von Metaboliten, wie zum Beispiel Oxalacetat oder 2-Oxoglutarat, die für das Zellwachstum benötigt werden [170].

Bis vor kurzem ist man davon ausgegangen, dass Cyanobakterien einen unvollständigen TCA-Zyklus besitzen, da die Gene, welche für den 2-Oxoglutarat-Dehydrogenase (OGDH)-Komplex kodieren, nicht in dessen Genomen annotiert sind. Dementsprechend besitzen bisher nahezu alle älteren, veröffentlichten Netzwerkrekonstruktionen von *Synechocystis* jeweils einen unvollständigen TCA-Zyklus und verwenden daher zusätzliche Reaktionen, um einen zyklischen Fluss zu ermöglichen. Ein Beispiel hierfür ist die Rekonstruktion, die von Shastri und Morgan [120] veröffentlicht wurde, welche den Glyoxylatzyklus benutzt, um den TCA-Zyklus zu schließen. Ebenso wie in den vorangegangenen Netzwerkkonstruktionen beziehungsweise in den ersten Versionen des Netzwerkes [93] wurde der TCA-Zyklus durch einen Bypass, den sogenannten γ -Aminobutyrat (GABA)-Shunt, geschlossen. Dabei läuft ein Fluss von 2-Oxoglutarat über Glutamat, γ -Aminobutyrat und Succinat-Semialdehyd, zu Succinat.

Erst vor kurzer Zeit wurde in *Synechococcus* sp. PCC 7002 eine direktere Möglichkeit von Zhang und Bryant gefunden, den TCA-Zyklus in Cyanobakterien zu schließen [142]. Viele Cyanobakterien, einschließlich *Synechocystis*, besitzen Gene für zwei Enzyme, welche den fehlenden OGDH-Komplex ersetzen können. Dazu gehören zum einen Gene für eine 2-Oxoglutarat-Decarboxylase (EC 4.1.1.71) (2OGDC) (sll1981) und zum anderen eine Succinat-Semialdehyd-Dehydrogenase (EC 1.2.1.16) (SSADH) (slr0370). Zusammen bilden diese beiden Enzyme eine direkte Verbindung, welche es den Cyanobakterien gestattet, den unvollständigen TCA-Zyklus zu schließen und einen zyklischen Fluss zu ermöglichen (siehe Abbildung 2.10 B).

Wie in Kapitel 2.2.4 dargelegt, zeigt die optimale Flussverteilung des photoautotrophen Wachstums, keinen zyklischen Fluss durch den TCA-Zyklus. Dieser dient haupt-

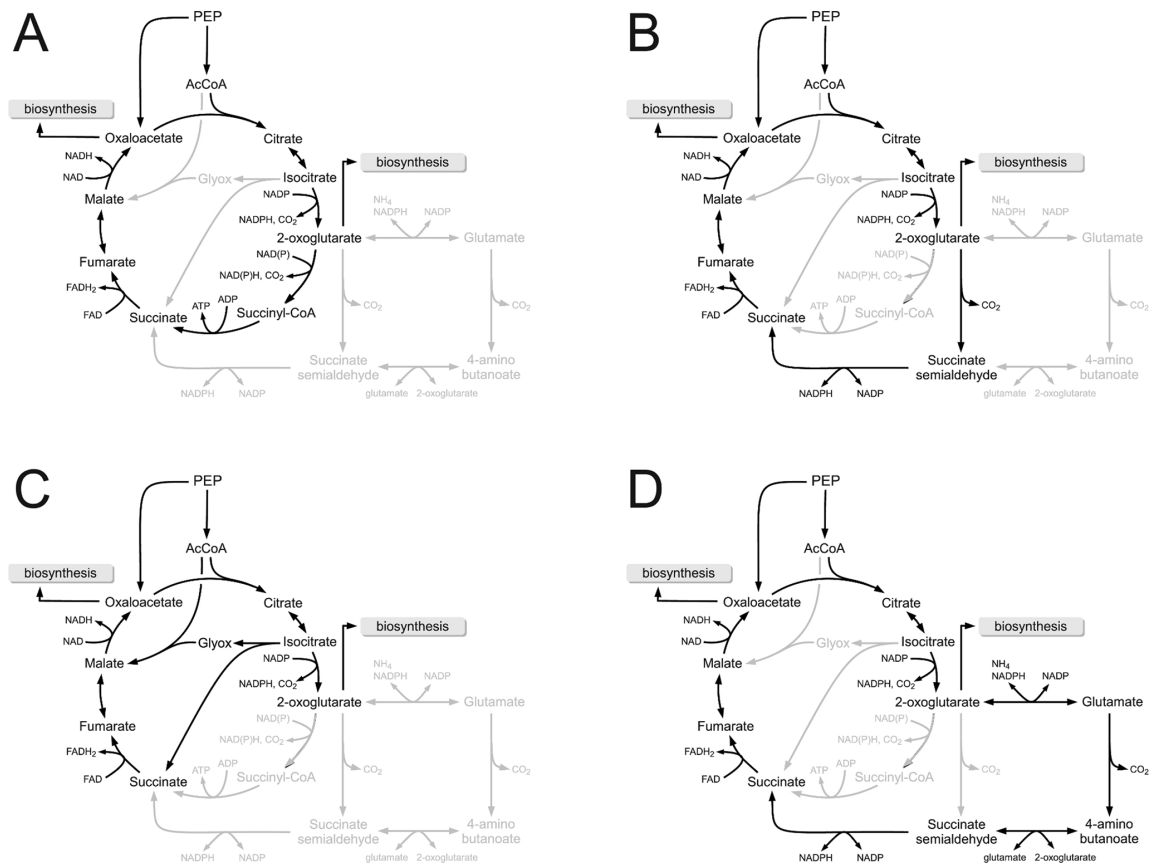


Abbildung 2.10.: Alternative Möglichkeiten des TCA-Zyklus. A) Der konventionell geschlossene, bakterielle TCA-Zyklus über einen OGDH-Komplex. B) Der TCA-Bypass von Zhang und Bryant [142] über eine 2-Oxoglutarat-Decarboxylase. C) Der Glyoxylat-Shunt, um den TCA-Zyklus zu überbrücken, verwendet unter anderem bei der Rekonstruktion von Shastri und Morgan [120]. D) Der GABA-Shunt wurde als erste Möglichkeit eingebaut, einen zyklischen Fluss über den TCA-Zyklus ohne Glyoxylat-Shunt zu ermöglichen [93]. Vom stöchiometrischen Aspekt her ist der konventionelle Zyklus A) am effektivsten (Abbildung aus Knoop et al. (2013) [94]).

sächlich dazu, Oxalacetat und 2-Oxoglutarat für die Synthese der Biomasse bereitzustellen und um Fumarat, welches während des Wachstums zum Beispiel bei der Purinsynthese entsteht, zurück in den Stoffwechsel zu führen. Demzufolge tragen die beiden neu entdeckten Enzyme bei der berechneten Flusslösung keinen Fluss. Nichtsdestotrotz ist das Nichtvorhandensein eines zyklischen Flusses durch den TCA-Zyklus, während des ausschließlich phototrophen Wachstums, nicht in vollständiger Übereinstimmung mit experimentell gemachten Beobachtungen. Zumindest besitzen *Synechococcus* sp. PCC 7002-Mutanten, bei denen die 2OGDC und SSADH fehlen beziehungsweise nicht funktionsfähig sind, eine geringere Wachstumsrate unter konstanten Lichtbedingungen und unter Hell-Dunkel-Bedingungen, als der Wildtyp [142].

Bei Abwesenheit von Licht ändert sich das vorhergesagte Flussmuster. In der Dunkelphase läuft typischerweise ein zyklischer Fluss durch den TCA-Zyklus ab, damit der respiratorische Stoffwechsel ablaufen kann, um die ATP-Synthese zu gewährleis-

ten (siehe Kapitel 2.2.6. Um die Rolle des zyklischen Flusses in Abwesenheit von Licht durch den TCA-Zyklus abzuschätzen, wird zwischen vier möglichen Szenarien unterschieden, welche in Cyanobakterien eine Schließung des TCA-Zyklus ermöglichen:

- A) einem konventionellen bakteriellen TCA Zyklus mit OGDH-Komplex, welcher nicht in Cyanobakterien annotiert ist
- B) einem cyanobakteriellen TCA-Zyklus mit Verwendung einer 2-Oxoglutarat-Decarboxylase und einer Succinat-Semialdehyd-Dehydrogenase
- C) einem Glyoxylatzyklus inklusive Isocitratlyase-Reaktion für einen zyklischen Fluss durch den TCA-Zyklus
- D) sowie einem unvollständigen TCA-Zyklus, der durch den GABA-Shunt geschlossen wird, um einen zyklischen Fluss zu ermöglichen

In der ersten veröffentlichten Rekonstruktion des Stoffwechselnetzwerkes von *Synechocystis* [93] (Abbildung 2.10 D) war ein Fluss durch den GABA-Shunt für die Zellatmung notwendig. Dabei ist die Abfolge der zugehörigen Reaktionen von der Stöchiometrie identisch zum relativ neu entdeckten Weg über die 2-Oxoglutarat-Decarboxylase von Zhang und Bryant [142] (Abbildung 2.10 B). Daher werden im Kontext einer FBA beide Wege mit der gleichen Effektivität genutzt.

Interessanterweise ist jedoch die Effektivität in Bezug auf die Biomasseproduktion für die beiden erst genannten Fälle niedriger als bei der Nutzung des konventionellen Zyklus über den OGDH-Komplex und somit eigentlich nur eine suboptimale Lösung für den respiratorischen Stoffwechsel. Nogales et al. [125] argumentieren, dass der GABA-Shunt trotzdem eine evolutionär günstigere Lösung wäre und begründen dies damit, dass ein festgesetzter Fluss durch den GABA-Shunt während des phototrophen Wachstums zu keiner Verringerung der Wachstumsrate führt im Vergleich zu einem festgesetzten Fluss durch den OGDH-Komplex. Gleichwohl bezogen auf den neuesten TCA-Weg nach Zhang und Bryant [142] kann dieser Tatbestand nicht unterstützt werden. Eine genauere Betrachtung dieses Sachverhalts ist in Kapitel 2.2.10 dargelegt.

Eine schlüssigere Erklärung für die Nutzung des stöchiometrisch ineffizienteren TCA-Weges bei Cyanobakterien über eine 2-Oxoglutarat-Decarboxylase und eine Succinat-Semialdehyd-Dehydrogenase anstelle des herkömmlichen OGDH-Komplexes, kann der Unterschied beider Wege in den Anforderungen in Bezug auf die Proteinsynthese sein: Der OGDH-Komplex ist eine hochentwickelte Maschinerie, die analog zum Pyruvatdehydrogenase (PDH)-Komplex durch drei Untereinheiten kodiert wird. In *Escherichia coli* besteht dieser Komplex aus einem 24-teiligen Kern aus E2-Proteinkomponenten, welche durch das *sucB*-Gen kodiert werden und einer unbekannten Anzahl der beiden anderen Komponenten in einem bis jetzt noch nicht vollständig geklärten Verhältnis [171, 172]. Hingegen werden die beiden alternativen Routen über den GABA-Shunt oder den Weg über die 2-Oxoglutarat-Decarboxylase von Zhang und Bryant [142] von strukturell einfacheren Enzymen kodiert. Ein Überblick über die Aminosäureanzahl für die jeweiligen Enzyme ist in Tabelle 2.4 zu sehen.

Angesichts der hohen strukturellen Komplexität des OGDH-Komplexes und der relativen Bedeutungslosigkeit des zyklischen Flusses durch den TCA-Zyklus für das phototrophe Wachstum von Cyanobakterien, könnte ein solcher Unterschied in den Proteinkosten ein Kompromiss zwischen der enzymatischen Effektivität und den Synthesekosten darstellen. Tatsächlich wird zunehmend beobachtet, dass die Maximierung des Stoffwechsels auf Effektivität und Biomassessyntheseeffizienz nicht notwendigerweise immer das universelle Prinzip ist [101, 173]. In diesem Kontext ist auch interessanterweise die Isocitratlyase zu nennen, welche nicht in *Synechocystis* vorhanden ist. Auch hier gilt, wenn diese in das Modell eingefügt wird, ergibt sich eine etwas höhere Ausbeute an Biomasse beim Wachstum bei Abwesenheit von Licht. Bei diesem Szenario dient das entstehende Succinat als Substrat für die SDH und erlaubt damit einen zyklischen Fluss durch den TCA-Zyklus für den respiratorischen Stoffwechsel.

Tabelle 2.4.: Enzymschritte mit der benötigten Anzahl (Länge) an Aminosäuren für eine Verbindung im TCA-Zyklus beziehungsweise dessen Bypass von 2-Oxoglutarat zu Succinat beziehungsweise Succinyl-CoA (Tabelle aus Knoop et al. (2013) [94]).

Organismus	Stoffwechselweg	Gen	EC-Nummer	Länge
<i>Escherichia coli</i> K-12	OGDH-Komplex	sucA	1.2.4.2	933
		sucB	2.3.1.61	405
		lpd	1.8.1.4	474
<i>Arabidopsis thaliana</i>	OGDH-Komplex	AT3G55410	1.2.4.2	1017
		AT5G65750	1.2.4.2	1025
		AT4G26910	2.3.1.61	464
		AT5G55070	2.3.1.61	464
		mtLPD1	1.8.1.4	507
		LPD1	1.8.1.4	623
		mtLPD2	1.8.1.4	507
<i>Chlamydomonas reinhardtii</i>	OGDH-Komplex	AT4G16155	1.8.1.4	630
		OGD1	1.2.4.2	1037
		OGD2	2.3.1.61	450
		DLD2	1.8.1.4	574
<i>Helicobacter pylori</i> 26695	OOR	GCSL	1.8.1.4	502
		oorA	1.2.7.3	375
		oorB	1.2.7.3	273
		oorC	1.2.7.3	186
<i>Synechocystis</i> sp. PCC 6803	2OGDC	oorD	1.2.7.3	113
		sll1981	4.1.1.71	550
	GABA-Shunt	gabD	1.2.1.16	454
		slr1022	2.6.1.19	429
		sll1641	4.1.1.15	467
		gabD	1.2.1.16	454

2.2.10. Der zyklische Fluss durch den TCA-Zyklus in der optimalen Flussverteilung

In der Literatur wird behauptet, dass ein Nebenweg über γ -Aminobutyrat (GABA-Shunt) gegenüber einem kompletten TCA-Zyklus einen evolutionären Vorteil bei autotrophen Bedingungen auf Kosten des reduzierten Wachstums bei heterotrophen Bedingungen besitzt [125]. Diese Hypothese wurde überprüft, indem ein Fluss durch einen potentiellen OGDH-Komplex sowie durch die Glutamat-Synthase erzwungen wurde, einmal unter autotrophen und einmal unter heterotrophen Bedingungen.

Ein Fluss durch den OGDH-Komplex führt zu einer sofortigen und starken Reduktion des Wachstums unter autotrophen Bedingungen, während ein erzwungener Fluss durch die Glutamat-Synthase (EC 1.4.1.14) (GLS) nur zu einer geringfügigen Reduktion des autotrophen Wachstums führt. Wie in Kapitel 2.2.4 beschrieben wird, ist jedoch für das autotrophe Wachstum kein geschlossener TCA-Zyklus erforderlich und ein festgesetzter Fluss durch einen OGDH-Komplex würde nur eine Umverteilung des optima-

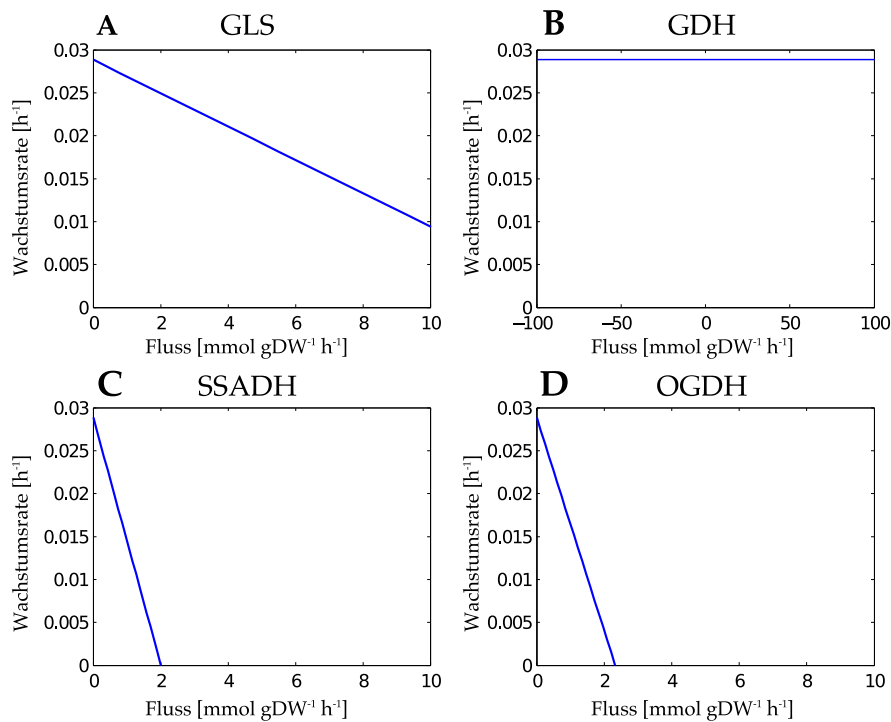


Abbildung 2.11.: Auswirkungen des TCA-Zyklus auf das optimale Wachstum unter phototrophen Bedingungen. Gezeigt sind hierbei die Wachstumsraten bei einem zunehmenden Fluss durch: A) GLS: die Glutamat-Synthase B) GDH: die Glutamat-Dehydrogenase C) SSADH: die Succinat-Semialdehyd-Dehydrogenase D) OGDH: den 2-Oxoglutarat-Dehydrogenase-Komplex. Während ein zunehmender Fluss durch die Glutamat-Dehydrogenase (GDH) bis zur unteren beziehungsweise oberen Flussgrenze keinen Einfluss auf die optimale Wachstumsrate hat, ist bei einem größeren Fluss durch die Glutamat-Synthase (GLS) ein Einfluss zu erkennen. Bei der Succinat-Semialdehyd-Dehydrogenase (SSADH) und der 2-Oxoglutarat-Dehydrogenase (OGDH) genügt schon ein niedriger Fluss, um die Wachstumsrate stark zu beeinträchtigen.

len Flusses erzwingen und somit eine unnötige metabolische Belastung darstellen. Auf der anderen Seite fließt auch unter autotrophen Bedingungen ein Fluss durch die GLS. Eine manuelle Festlegung des Flusses durch die GLS hat daher keine großen unmittelbaren Auswirkungen auf das Wachstum. Wenn aber anstatt eines Flusses durch die GLS ein Fluss durch die Succinat-Semialdehyd-Dehydrogenase erzwungen wird, welche auch einen Schritt des GABA-Shunts darstellt, erhält man ähnliche Ergebnisse wie für den Fluss durch den OGDH-Komplex (siehe Abbildung 2.11). Die evolutionäre Bedeutung und die jeweiligen möglichen Vorteile des GABA-Shunts gegenüber dem OGDH-Komplex und dem neu entdeckten TCA-Nebenweg (Bypass) von Zhang und Bryant (2011) [142], bleiben demnach noch offen.

2.2.11. Zeitliche Koordination des phototrophen Stoffwechsels

In den meisten natürlichen Lebensräumen der Cyanobakterien müssen sie ihren Stoffwechsel an unterschiedliche Lichtbedingungen anpassen, die während eines Tagesverlaufs auftreten. Diese Maßgaben haben einen signifikanten Einfluss auf den Organismus und bringen ihn dazu, seinen Stoffwechsel innerhalb des metabolischen Netzwerkes zu reorganisieren und zu verändern. Dementsprechend sind Cyanobakterien die einzigen bis jetzt bekannten Prokaryoten, die eine endogene circadiane Uhr besitzen, die als intrazellulärer Zeitgeber fungiert [174].

Um das zyklische Verhalten von Cyanobakterien genauer zu erforschen, gab es über die letzten Jahre viele Studien, zum Beispiel mit Hochdurchsatzdaten des Stoffwechsels [145, 175–178]. Allerdings konzentrieren sich alle aktuellen großen Netzwerkrekonstruktionen ausschließlich auf ein phototrophes Wachstum unter konstantem Lichteinfall oder auf ein heterotrophes Wachstum. In dieser Arbeit wird versucht, diese Lücke durch eine Analyse der zeitlichen Koordination des cyanobakteriellen Stoffwechsels zu schließen, indem ein voller Tagesablauf des phototrophen Stoffwechsels simuliert wurde. Das Einbeziehen und Modellieren von zirkadianen Änderungen in ein genomskaliges Modell des Stoffwechsels ist keine triviale Aufgabe. Um dies zu lösen, stehen im Allgemeinen zwei Ansätze zur Verfügung:

- Der Bottom-up-Ansatz:

Hierbei können große Mengen von Gentranskriptionsdaten oder Proteinexpressionsdaten verwendet werden, um die Präsenz von Proteinen beziehungsweise enzymatischen Umwandlungen zu bestimmten Zeitpunkten zu ermitteln. Dabei ist jedoch zu beachten, dass das Vorhandensein von Gentranskripten oder Proteinen nicht notwendigerweise mit dem metabolischen Fluss korreliert und es wurden schon häufig widersprüchliche Expressionswerte für einzelne Wege oder Isoenzyme gemessen. Eine aktuelle Analyse, in der gepaarte mRNA- und Proteindaten für Tag-Nacht-synchronisierte Kulturen des Cyanobakteriums *Prochlorococcus* MED4 ausgewertet wurden, zeigte nur eine geringe Korrelation zwischen mRNA- und Proteinvorkommen. Zudem wurden nur geringe Änderungen in der relativen Abundanz aller Enzyme über den gesamten zeitlichen Verlauf beobachtet [179]. Daher wird vermutet, dass eine direkte Integration von Gentranskriptionsdaten um metabolische Flüsse zu regeln beziehungsweise einzuschränken, wie es bereits für heterotrophe Bakterien angewendet wurde [180],

keine geeignete Strategie ist, um den periodischen Tagesverlauf des cyanobakteriellen Stoffwechsels zu beschreiben.

- Der Top-down-Ansatz:

Eine alternative Strategie, welche in dieser Arbeit verfolgt wurde, ist daher ein Top-down-Ansatz, bei dem die Zielfunktion der Zelle abhängig von der Zeit und den Veränderung der Lichtverfügbarkeit definiert wird. Speziell hierfür wurde ein kürzlich erworbener Datensatz vom zyklischen Verhalten der Gentranskripte von *Synechocystis* sp. PCC 6803 [181] verwendet (bereitgestellt von Robert Lehmann), welcher Einblicke in die zeitliche Koordination des cyanobakteriellen Stoffwechsels gewährt. Hierbei wurde der Transkriptdatensatz, entgegen dem Einsatz bei einem Bottom-up-Ansatz, nicht direkt dazu verwendet, um bestimmte Enzymaktivitäten und damit verbundene Reaktionsraten festzulegen oder einzugrenzen, sondern um einen umfassenden Überblick über die zeitliche Koordination ganzer Enzymgruppen im Stoffwechsel zu erhalten.

Die sich hieraus ergebenden Expressionsmuster von Stoffwechselenzymen sind in Abbildung 2.12 gezeigt und liefern wichtige Hinweise auf die zeitliche Koordination des phototrophen Stoffwechsels. Besonders fällt auf, dass in *Synechocystis* im Gegensatz zu einigen diazotrophen Stämmen [175], die Mehrheit der oszillierenden Gene ihr Transkriptionsmaximum während des Tages erreichen, was auf eine stark reduzierte Expressionsaktivität während der Nacht schließen lässt. Von den Transkripten, deren Expressionsrate während der Nacht am höchsten ist, können die meisten Transportvorgängen zugeordnet werden oder auch zum Teil dem TCA-Zyklus. Hierzu zählt insbesondere das Transkript, welches dem Gen entspricht (sll1981), das von Zhang und Bryant [142] für den TCA-Bypass identifiziert wurde. Aufgrund der Tatsache, dass viele Transportprozesse für die Aufnahme von wachstums-limitierenden Mikronährstoffen, wie Eisen oder Mangan, zuständig sind, gibt es wenig Grund zu der Annahme, dass der Transport exklusiv auf den Tag beschränkt ist.

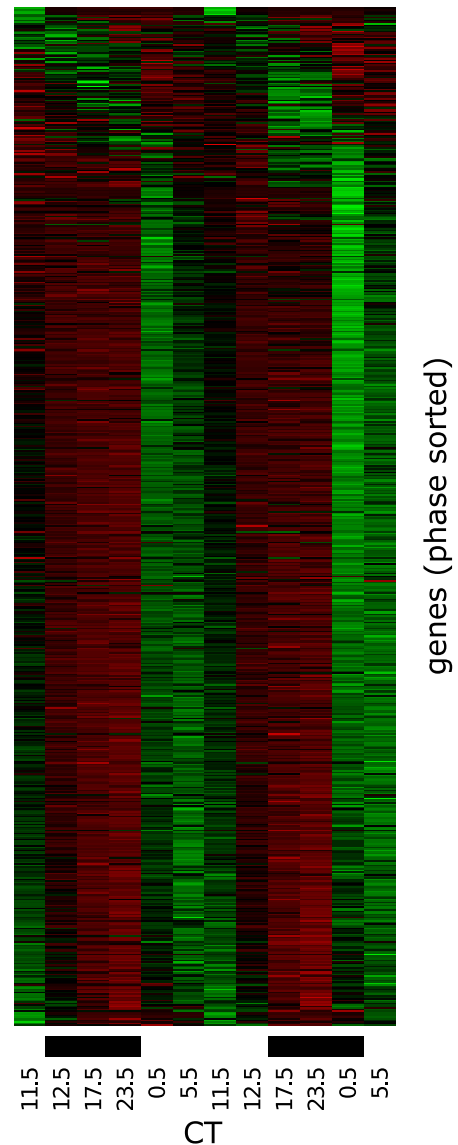


Abbildung 2.12.: Phasensortierte Expressionsprofile von metabolischen Genen in *Synechocystis* als eine Funktion der circadianen Zeit (CT). Es sind zwei unabhängige Replikate nacheinander gezeigt. Die große Mehrheit der Transkripte hat ihr Expressionsmaximum während der Lichtphase (Abbildung aus Knoop et al. (2013) [94]).

Die Einblicke, die man durch Expressionsdaten erhält, können zusätzlich durch physiologische Daten ergänzt werden. So existieren zum Beispiel für *Synechocystis* oder anderen Cyanobakterienstämme wie *Cyanothece* ATCC 51142 zeitliche Messungen beziehungsweise Abschätzungen der Biomasse und des Chlorophyllgehalts durch Bestimmung der optischen Dichte bei verschiedenen Wellenlängen. Es konnte gezeigt werden, dass die Chlorophyllkonzentration vom frühen Morgen bis zum Mittag stark ansteigt und anschließend konstant bleibt [176]. Demgegenüber erhöht sich die Biomasse erst signifikant nach Ende der Chlorophyllakkumulation, was sich durch eine entsprechende Zunahme von Speicherverbindungen wie Glykogen erklären lässt [176]. Ferner wurde für *Synechococcus elongatus* PCC 7942 gezeigt, dass trotz einer rhythmischen Zellteilung die DNA-Synthese immer mit einer konstanten Rate erfolgt, das heißt, es existiert keine Koppelung zwischen der DNA-Synthese und der Zellteilung [182].

Basierend auf diesen empirischen Beobachtungen, wurde eine zeitaufgelöste Biomassezielfunktion (*Biomass Objective Function*) entwickelt, um einen vollen Tagesablauf des phototrophen Stoffwechsels von *Synechocystis* zu simulieren. Statt der Benutzung einer einzelnen BOF wurden die unterschiedlichen Biomassekomponenten nach ihrer Synthesereihenfolge wie folgt aufgeschlüsselt: Es wurde davon ausgegangen, dass die Aufnahmerate von Mikronährstoffen beziehungsweise anorganischen Ionen über den ganzen Tagesverlauf gleich bleibt. Auch die DNA-Synthese wird über den gesamten Zeitraum konstant gesetzt. Ebenso wird eine gleichbleibende ATP-Verbrauchsrate für den basalen ATP-Verbrauch (*Maintenance*) über den Tag angenommen. Alle anderen Biomassekomponenten werden durch einen Faktor in der BOF beschrieben, die nach Lichtverfügbarkeit maximiert wird.

Um die Ergebnisse der Pigmentmessungen [176] abzubilden, wurde der Faktor für die Pigmentsynthese zwei Stunden vor Sonnenaufgang erhöht und ab Mittag wieder vermindert. Im Gegensatz dazu wurde der Faktor für die Synthese der Speicherstoffe beziehungsweise von Glykogen erst nach der Mitte des Tages erhöht. Zusätzlich wurde ein ATP-Verbrauch für die Proteinsynthese in die BOF eingebracht, der einen erhöhten Bedarf an ATP während des Wachstums widerspiegeln soll. Ähnlich zu den Simulationen bei Abwesenheit von Licht (siehe Kapitel 2.2.6), wird eine untere Grenze des Flusses für die Zellatmung auch während der Lichtperioden gesetzt, so dass eine restliche respiratorische Aktivität auch während des Tages gewährleistet wird. Die Entwicklung von radikalen Sauerstoffspezies und die Mehler-Reaktion werden abhängig zur Lichtintensität gesetzt. Der Stoffwechsel während der Nacht wird auf den Glykogenabbau, welcher zur DNA-Synthese, zur Aufnahme von Nährstoffen und anorganischen Ionen sowieso zur Zellatmung und zur Erfüllung des basalen ATP-Verbrauchs (*Maintenance*) dient, beschränkt. Für die Biomassesynthese wurde das Prinzip der dynamischen FBA [183] genutzt.

Die Ergebnisse eines vollen Tagesablaufs des phototrophen Stoffwechsels sind in Abbildung 2.13 gezeigt. Dabei sind komplexe Übergänge in den metabolischen Flüssen über den gesamten Tagesverlauf von 24 Stunden zu beobachten. Angefangen vom Metabolismus in der Nacht, der von dem respiratorischen Stoffwechsel dominiert wird,

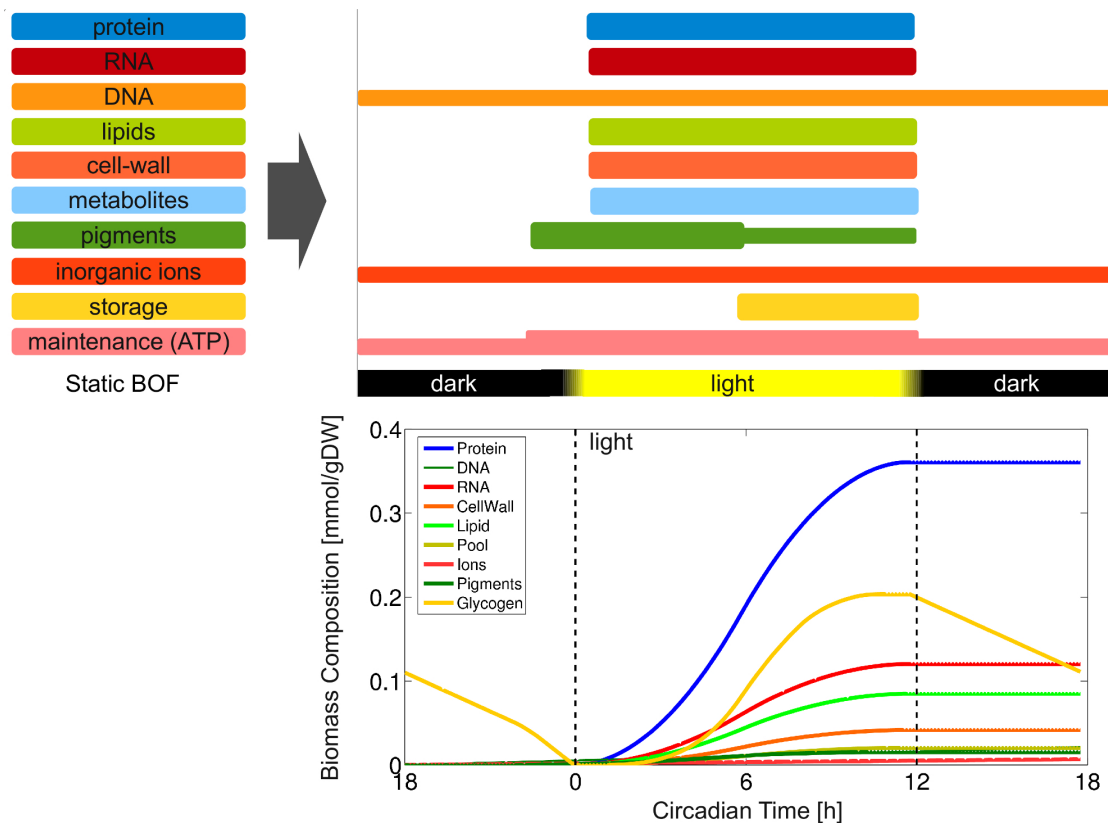


Abbildung 2.13.: Zeitabhängige Biomassezielfunktion. Die obere Graphik zeigt die Aufteilung der Biomassefunktion über einen vollen Tagesgang. Der untere Graph zeigt die Simulation der Synthesen und Akkumulationen der einzelnen Biomassekomponenten in Abhängigkeit zur circadianen Zeit (CT). CT = 0 h markiert den Sonnenaufgang (Abbildung aus Knoop et al. (2013) [94]).

bis hin zum optimierten Zellwachstum im Laufe des Tages. Wie vorher definiert, erhöht sich die Konzentration der Pigmente in den frühen Morgenstunden und bleibt nach der Mitte des Tages annähernd konstant. Der Glykogengehalt nimmt während der zweiten Hälfte des Tages stark zu und wird während der Nacht als Kohlenstoff- und Energiequelle verwendet.

Für einige ausgewählte metabolische Reaktionen sind die zeitabhängige Flussraten in der Abbildung 2.14 dargestellt. Gezeigt sind dabei:

- A) die Netto-Sauerstoffaufnahme, die bei der Abwesenheit von Licht positiv ist und am Tage proportional zur Lichtverfügbarkeit verläuft
- B) die Phosphoglyceratkinase, welche zuerst in der Nacht einen geringen negativen Fluss gemäß des Glykogenabbaus und später tagsüber einen hohen positiven Fluss entsprechend der Regeneration des Calvin-Benson-Zyklus aufweist.
- C) die Interkonversion von Glucose-1-Phosphat zu Glucose-6-Phosphat läuft zunächst entsprechend des Glykogenabbaus im Dunkeln Richtung der Glykolyse ab, also zu

Glucose-6-Phosphat (negative Flusswerte), während des Tages zum Aufbau von Glykogen und der Synthese von Lipiden aber in die entgegengesetzte Richtung zu Glucose-1-Phosphat (positive Flusswerte)

- D) der Fluss durch die RuBisCO-Reaktion, die ein Maß für die photosynthetische Aktivität und die Verfügbarkeit von Energie ist
- E) die CTP-Kinase-Reaktion von CTP zu CDP und umgekehrt, welche durch die DNA-Synthese und durch die Pigmentsynthese beeinflusst wird
- F) sowie die Succinat-Semialdehyd-Dehydrogenase, welche den TCA-Zyklus schließt und dadurch einen positiven Fluss in der Nacht zeigt, während diese am Tag inaktiv ist

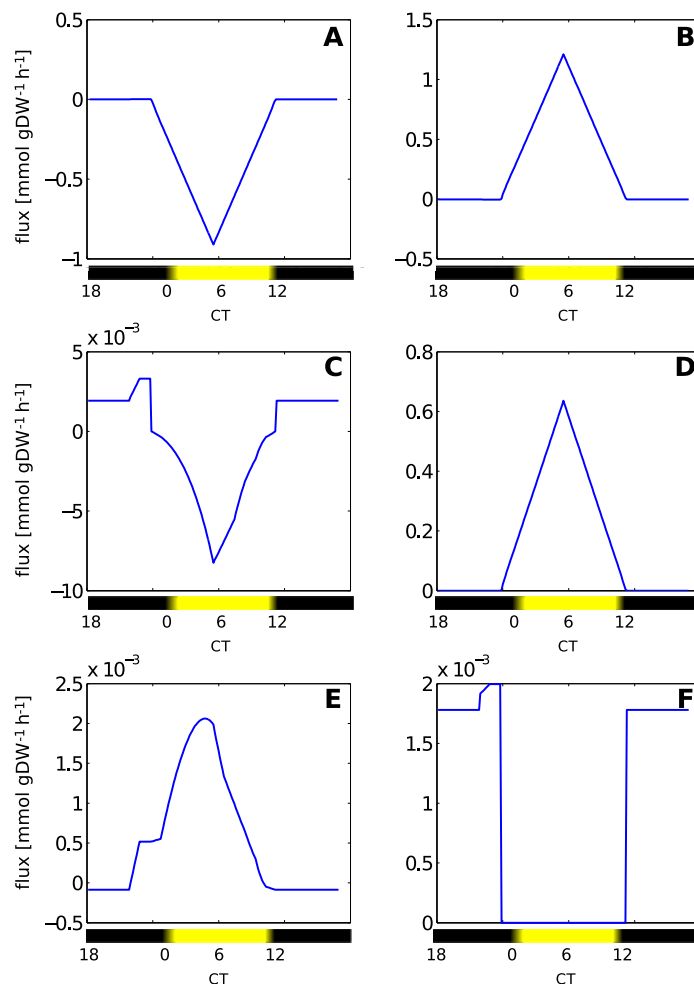


Abbildung 2.14.: Zeitlicher Verlauf des metabolischen Flusses einiger ausgewählter Enzyme während eines vollen Tagesganges. Dargestellt sind: (A) Netto-Sauerstoffrate (B) Phosphoglyceratkinase (C) Phosphoglucomutase (D) RuBisCO (E) CDP-Kinase (F) Succinat-Semialdehyd-Dehydrogenase (Abbildung aus Knoop et al. (2013) [94]).

2.3. Schlussfolgerungen

In der heutigen Zeit wird die Analyse der Struktur und Funktion von metabolischen Netzwerken zunehmend durch rechnerische Ansätze unterstützt, welche eine systematische Sichtweise auf die möglichen Interkonversionsrouten innerhalb des Netzwerkes gewähren. Für eine solche Analyse ist eine umfassende Rekonstruktion, die alle enzymatischen Reaktionen, die in einer Zelle beziehungsweise einem Organismus ablaufen, umfasst, eine wichtige Voraussetzung. In dieser Hinsicht wurde in dieser Arbeit eine manuell kuratierte Rekonstruktion des metabolischen Netzwerkes von *Synechocystis* sp. PCC 6803 erstellt, welcher ein bekannter einzelzelliger Modellorganismus für den phototrophen Stoffwechsel ist. Die Rekonstruktion eines metabolischen Netzwerkes basiert in der Regel auf der annotierten Genomsequenz des spezifischen Organismus, kombiniert mit verschiedenen Reaktions- und Stoffwechselrepertoires beziehungsweise Datenbanken. Allerdings liegt der Fokus der meisten Datenbanken und Repositorien in erster Linie auf einzelnen Ausschnitten des Stoffwechsels und sie betrachten deshalb nicht das resultierende Netzwerk als eine ganze gesamte funktionale Einheit. Im Gegensatz dazu zielt die Netzwerkrekonstruktion insbesondere darauf, die vielfältigen Anforderungen, die eine bloße Liste von Reaktionen gegenüber einem funktionalen Netzwerk unterscheiden, mit einzubeziehen. Solche Anforderungen umfassen die Fähigkeit des Netzwerkes, alle essentiellen Zwischenprodukte und Vorstufenmetaboliten, die für die Produktion von Biomasse benötigt werden, zu synthetisieren, eine Möglichkeit aus dem System heraus Energie (ATP) für zum Beispiel den basalen ATP-Verbrauch (*Maintenance*) zu gewinnen, das Redox-Gleichgewicht zu erhalten und auch in der Lage zu sein, alle Metaboliten und Cofaktoren des Netzwerkes wieder zu regenerieren. Falls eine oder mehrere dieser Anforderungen nicht von der Netzwerkrekonstruktion erfüllt werden, weist dieses auf Fehler oder Lücken in der Stöchiometrie des Netzwerkes hin. Dies hätte eine entsprechende erneute Prüfung des Netzwerkes zur Folge. Somit kann die Rekonstruktion metabolischer Netzwerke, welche den zellulären Stoffwechsel als hoch integratives Netzwerk ansieht, als Vorstufe zur metabolischen Modellierung gesehen werden.

Trotz der bedeutenden Leistungen der letzten Jahre durch mehrere Dutzend Rekonstruktionen von unterschiedlichen Organismen, befindet sich die systematische Rekonstruktion genomskaliger metabolischer Netzwerke von phototrophen Organismen noch in ihrem Anfangsstadium. Während bereits viele Studien und Analysen für Algorithmen zur automatisierten Lückenschließung in Netzwerken existieren, sind diese Algorithmen in der Regel nur an hoch idealisierten Beispielen veranschaulicht, wie zum Beispiel die Wiederherstellung eines zuvor bekannten kompletten Netzwerkes nach manuellem Entfernen einzelner Reaktionen. Aus diesem Grund basiert bis jetzt eine qualitativ hochwertige Rekonstruktion überwiegend auf einer umfangreichen manuellen Kuration, die ihre Informationen aus der biochemischen Primärliteratur bezieht.

Ein anhaltendes Problem sind die vielfältig verwendeten Namenskonventionen in verschiedenen Datenbanken und bestehenden Rekonstruktionen sowie die manchmal

inkonsistente oder falsche Annotation von Enzymen, die oft nur auf Sequenzvergleichen basiert. Eine zeitnahe Lösung für dieses Problem könnte die Entwicklung von Standards beziehungsweise die Durchsetzung von Normen und Konventionen in wissenschaftlichen Veröffentlichungen sein [184]. Andere Probleme rühren meist von spezifischen Variationen her und sind weniger leicht zu überwinden. Zum Beispiel wurde kürzlich gezeigt, dass selbst innerhalb eines solchen grundlegenden und gut analysierten Stoffwechselweges, wie dem zentralen Kohlenstoffstoffwechsel in *Escherichia coli*, neue und unbekannte Reaktionen immer noch entdeckt werden können [185]. In dieser Hinsicht kann jetzt nicht erwartet werden, dass die in dieser Arbeit vorgestellten Rekonstruktionen der metabolischen Netzwerke von *Synechocystis* sp. PCC 6803 oder *Synechococcus elongatus* PCC 7942 (siehe Kapitel 3) eine fehlerfreie Darstellung des *in vivo*-Zustandes widerspiegeln. Vielmehr liegt die Absicht dieser Arbeit auch darin, neue Diskussionen über mögliche Inkonsistenzen oder Schwachpunkte des Wissensstandes über den cyanobakteriellen Stoffwechsel anzuregen und weitere Ideen zu fördern, die diese Unklarheiten auflösen oder beheben. Insbesondere bietet eine qualitativ hochwertige Netzwerkrekonstruktion, auch wenn sie zum Teil fehlerhaft oder unvollständig ist, enorme Möglichkeiten, die Funktionsweise des Stoffwechsels zu erforschen und zu verstehen.

Die in dieser Arbeit vorgestellte genomskalige Rekonstruktion des Cyanobakteriums *Synechocystis* sp. PCC 6803 beinhaltet das Wissen und die Informationen, welche bis zum aktuellen Zeitpunkt für dieses Cyanobakterium verfügbar waren. Integriert sind dabei neuartige Ergebnisse bezüglich des cyanobakteriellen TCA-Zyklus, über den vermeintlich vorhandenen Glyoxylatzyklus sowie die Rolle der Photorespiration im Zusammenhang mit dem Zellwachstum. Das Modell beinhaltet verschiedene Aspekte, die spezifisch für den phototrophen Stoffwechsel sind, wie unter anderem eine lichtabhängige Bildung von reaktiven Sauerstoffspezies. Zusätzlich zu dem Modell selber, welches im verbreiteten SBML-Format vorliegt und somit für weitere Modellierungszwecke leicht zur Verfügung steht, wurde eine detaillierte grafische Übersicht des metabolischen Netzwerkes von *Synechocystis* angefertigt, welche weitere Diskussionen und die Interpretation von Ergebnissen erleichtert. Der Fokus lag hierbei auf der Analyse des phototrophen Wachstums des Organismus, insbesondere wurden auch die Konsequenzen und Einflüsse von einer abweichenden Topologie des Netzwerkes auf das Wachstum beziehungsweise den gesamten Stoffwechsel untersucht.

In der Tat bleiben trotz der zahlreichen Rekonstruktionen in jüngster Zeit und der umfangreichen biochemischen Literatur, die für *Synechocystis* zur Verfügung steht, einige wichtige Reaktionsschritte unklar. In dieser Arbeit wurden mehrere alternative Möglichkeiten zur Schließung des TCA-Zyklus bei Cyanobakterien analysiert und der Biomasseertrag beziehungsweise das Wachstum für die verschiedenen Szenarien untereinander verglichen. Ein evolutionärer Vorteil des GABA-Shunts gegenüber dem traditionellen OGDH-Komplex, wie Nogales et al. [125] argumentieren, konnte nicht bestätigt werden. Der kürzlich entdeckte TCA-Bypass von Zhang und Bryant [142] sowie auch der GABA-Shunt, erfordern wesentlich kleinere Investitionen bezüglich der Enzymsynthese im Vergleich zum konventionellen OGDH-Komplex und könnte

daher evolutionär gesehen von Vorteil sein für Organismen, die in erster Linie vom phototrophen Wachstum abhängig sind.

Um die verschiedenen in der Literatur verfügbaren Rekonstruktionen zu bewerten, wurde experimentell von Marianne Gründel und Yvonne Zilliges getestet, ob der vermeintliche Glyoxylatzyklus, der auf Grund von biochemischen Beobachtungen in mehreren Rekonstruktionen implementiert wurde, vorhanden ist. Im Endeffekt konnte die Existenz einer enzymatischen Aktivität der Isocitratlyase unter den getesteten Bedingungen weder bestätigt werden, noch konnte gezeigt werden, dass *Synechocystis* in der Lage ist, auf Acetat in Gegenwart von DCMU zu wachsen. DCMU inhibiert die Oxidation von Wasser am Photosystem II und verhindert demzufolge den linearen Elektronentransport, aber nicht den zyklischen Elektronentransport am Photosystem I und die ATP-Generation. Beide Sachverhalte weisen stark darauf hin, dass kein funktionaler Glyoxylatzyklus in *Synechocystis* vorhanden ist.

Die meisten aktuellen Rekonstruktionen werden anhand der Flussbilanzanalyse, die es erlaubt, optimale Flussmuster, unter Einbeziehung einer speziellen Zielfunktion, vorherzusagen und zu bewerten, analysiert. Innerhalb dieser Arbeit half die FBA unter anderem dabei, Schlüsse über das phototrophe Wachstum und dessen optimalen Reaktionsflüsse zu ziehen sowie essentielle Gene und Synthesewege des photosynthetischen Stoffwechsels zu identifizieren. Zudem wurde gezeigt, wie sich der Stoffwechsel beim Übergang von Licht- zu Dunkelphasen beziehungsweise bei tagestypischen wechselnden Hell-Dunkel-Zyklen reorganisiert, was ein spezifisches Merkmal für phototrophe Organismen ist. Bisher zielten die in der Literatur verfügbaren Rekonstruktionen auf eine Analyse und Auswertung des hetero, mixo- und phototrophen Wachstums unter konstanten Lichtbedingungen, mit wenig Bezug zu den tatsächlichen Umgebungsbedingungen, denen der Organismus *in vivo* ausgesetzt ist. Dabei helfen Genexpressionsdaten und physiologische Eigenschaften, die es ermöglichen, grundlegende tageszeitpunktabhängige Abläufe im Stoffwechsel des Organismus zu beschreiben. Natürlich sind weitere ähnliche Ansätze erforderlich, um ein besseres Verständnis der Prinzipien und Abstimmungen während des phototrophen Wachstums zu erzielen. Zudem wirft die Auswertung der Simulation eines gesamten Tagesverlaufs mehrere Fragen in Bezug auf den cyanobakteriellen Stoffwechsel auf, die eine genauere Betrachtung und Aufmerksamkeit in der Zukunft verdienen.

Insbesondere der Stoffwechsel bei Abwesenheit von Licht, zu dem auch das Wechselspiel zwischen oxygener Photosynthese und aerober Atmung, die in einem einzigen Kompartiment der Zelle ablaufen, gehört, sind bis jetzt noch unzureichend erforscht. Letzteres unterstreicht die Verbindung und das Zusammenarbeiten zwischen den zellulären Stoffwechselwegen als ein System von Versorgung und Nachfrage, das durch den metabolischen Zustand der Zelle beeinflusst wird. Hierbei sind zahlreiche charakteristische Umschaltunkte auszumachen, die verdeutlichen, dass die Effizienz von zellulären Stoffwechselreaktionen keine intrinsische Eigenschaft des jeweiligen Stoffwechselweges ist, sondern von dem aktuellen Zustand und der metabolischen Nachfrage des gesamten Stoffwechsels abhängt.

Die Diversität des cyanobakteriellen Stoffwechsels

Dieses Kapitel basiert auf der Veröffentlichung:

[130] C. Beck, H. Knoop, I. M. Axmann und R. Steuer (2012). The diversity of cyanobacterial metabolism: genome analysis of multiple phototrophic microorganisms. *BMC Genomics*, 13(1):56

Vorwort

Im vorangegangenen Kapitel dieser Arbeit wurde im Detail auf den Stoffwechsel des Cyanobakteriums *Synechocystis* sp. PCC 6803 eingegangen und dieser mithilfe der *Constraint*-basierten Flussbilanzanalyse genauer untersucht.

Das nun anschließende Kapitel beschreibt die genaue Betrachtung und den Vergleich von 15 weiteren Cyanobakterien (16 Stämme inklusive *Synechocystis* sp. PCC 6803) auf Genomebene, wobei besonders Übereinstimmungen und Unterschiede auf Ebene des Stoffwechsels hervorgehoben werden. Die Beantwortung der Frage nach den Gemeinsamkeiten dieser Organismen des gleichen Phylums wird mithilfe der Definition von Clustern von vermutlichen orthologen Genen (CLOGs) eingeleitet. Diese ermöglichen zwischen Kern-Genen, gemeinsamen und einzigartigen Genen zu unterscheiden und somit Rückschlüsse über das Kern- und das Pan-Genom der Cyanobakterien zu ziehen. Gemeinsamkeiten unter den Cyanobakterien werden im Detail in Bezug auf den Zentralstoffwechsel und die Speicherstoffe genauer betrachtet, also welche Enzyme konserviert sind und wo Variabilität unter den einzelnen Cyanobakterien existiert.

Die Methoden und Ergebnisse, die in diesem Kapitel vorgestellt werden, wurden in Kooperation mit Christan Beck erarbeitet. Die Analysen, die im ersten Abschnitt dieses Kapitels aufgeführt werden, die Definition und Berechnung der CLOGs (Kapitel 3.1.2 bzw. 3.2.1) sowie die Anreicherung der GO-Annotation (Kapitel 3.1.3 bzw. 3.2.2) und die phylogenetische Analyse (Kapitel 3.1.4 bzw. 3.2.3), wurden von ihm durchgeführt.

3.1. Material und Methoden

3.1.1. Auswahl der Cyanobakterienstämme

Für die vergleichende Analyse der Diversität unter den Cyanobakterien wurden folgende 16 Stämmen ausgewählt:

- *Acaryochloris marina* MBIC 11017 (Aca11017)
- *Cyanothece* ATCC 51142 (Cyn51142)
- *Cyanothece* PCC 8801 (Cyn8801)
- *Gloeobacter violaceus* PCC 7421 (Glo7421)
- *Microcystis aeruginosa* NIES-843 (Mic843)
- *Nostoc* sp. PCC 7120 (Nos7120)
- *Prochlorococcus marinus* MED4 (ProMED4)
- *Prochlorococcus marinus* MIT 9211 (Pro9211)
- *Prochlorococcus marinus* MIT 9215 (Pro9215)
- *Synechococcus* sp. JA-2-3B'a(2-13) (SycJA23)
- *Synechococcus* sp. PCC 7002 (Syc7002)
- *Synechococcus* sp. WH7803 (Syc7803)
- *Synechococcus elongatus* PCC 7942 (Syc7942)
- *Synechocystis* sp. PCC 6803 (Syn6803)
- *Thermosynechococcus elongatus* BP-1 (ThermoBP1)
- *Trichodesmium erythraeum* IMS101 (Trich101)

Eine kurze Einleitung zu den Stämmen ist in Kapitel 1 gegeben. Die jeweiligen chromosomalen Genomsequenzen wurden im August 2010 aus der GenBank-Datenbank (<http://www.ncbi.nlm.nih.gov/GenBank>) [186] extrahiert. Plasmidsequenzen wurden für diese Analyse nicht weiter berücksichtigt, da nicht vollständig aufgeklärt ist, welche Gene auf diesen DNA-Fragmenten kodiert sind. Außerdem könnte die Einbeziehung solcher Sequenzen zu einer möglichen Verzerrung der Ergebnisse zugunsten von Genomen mit Plasmiden, welche mehrere Gene enthalten, die über einen möglichen horizontalen Gentransfer gewonnen wurden, im Vergleich zu Stämmen ohne Plasmide führen [187].

Die Auswahl der Stämme beruhte auf den Ergebnissen, die von Gupta und Matthews (2010) [188] beschrieben wurden. Basierend auf einem phylogenetischen Baum, der aus den Sequenzen von 44 konservierten Proteinen erstellt wurde, war es das Ziel, möglichst Stämme auszuwählen, die die Vielfalt der Cyanobakterien repräsentativ widerspiegeln. Ein Überblick über die Eigenschaften der ausgewählten Stämme ist in Tabelle 3.1 gegeben.

Tabelle 3.1.: Eine Zusammenfassung der Eigenschaften der 16 untersuchten Cyanobakterienstämme. Aufgeführt ist die spezifische Abkürzung des Stammes, der RuBisCO-Typ [162], die Genomgröße in Mb, der G+C Gehalt, die Anzahl annotierter Gene, der prozentuale Anteil kodierender DNA gemäß der IMG-Datenbank [189], die Fähigkeit der Stickstofffixierung, das Habitat und die vorherrschende Zellform. In der Kategorie des Habitats steht (M) für einen marinen Stamm, (S) für einen Süßwasserstamm und (T) für einen thermophilen Stamm. Bei der Zellform wird zwischen Einzelzellen (S) und filamentösen, mehrzelligen Zellformen (F) unterschieden. Die Unterteilung der Sektion des Stamms erfolgte gemäß Rippka et al. (1979) [190] (Tabelle aus Beck et al. (2012) [130]).

	Abkürz.	Typ	Genom- größe (Mb)	G+C	Gene	DNA kod. (%)	N ₂ - Fix.	Habitat	Zell- form	Sektion
<i>Acaryochloris marina</i> MBIC11017	Aca11017	β	8,36	46,96	8488	83,26	-	M	S	I
<i>Cyanothece</i> sp. ATCC 51142	Cyn51142	β	5,46	37,94	5354	86,80	•	M	S	I
<i>Cyanothece</i> sp. PCC 8801	Cyn8801	β	4,79	39,76	4615	84,85	•	S	S	I
<i>Gloeobacter violaceus</i> PCC 7421	Glo7421	β	4,66	62,00	4490	89,36	-	S	S	I
<i>Microcystis aeruginosa</i> NIES-843	Mic843	β	5,84	42,33	6360	81,43	-	S	S	I
<i>Nostoc</i> sp. PCC 7120	Nos7120	β	7,21	41,27	6222	82,50	•	S	F	IV
<i>Prochlorococcus marinus</i> MED4	ProMED4	α	1,66	30,80	1766	88,42	-	M	S	I
<i>Prochlorococcus marinus</i> MIT 9211	Pro9211	α	1,69	38,01	1901	90,12	-	M	S	I
<i>Prochlorococcus marinus</i> MIT 9215	Pro9215	α	1,74	31,15	2059	89,62	-	M	S	I
<i>Synechococcus</i> sp. JA-2-3B'a(2-13)	SycJA23	β	3,05	58,45	2947	85,48	•	S/T	S	I
<i>Synechococcus</i> sp. PCC 7002	Syc7002	β	3,41	49,19	3237	87,64	-	M	S	I
<i>Synechococcus</i> sp. WH7803	Syc7803	α	2,37	60,24	2591	93,39	-	M	S	I
<i>Synechococcus elongatus</i> PCC 7942	Syc7942	β	2,80	55,43	2719	89,21	-	S	S	I
<i>Synechocystis</i> sp. PCC 6803	Syn6803	β	3,57	47,37	3628	86,74	-	S	S	I
<i>Thermosynechococcus elongatus</i> BP-1	ThermoBP1	β	2,59	53,92	2555	89,99	-	S/T	S	I
<i>Trichodesmium erythraeum</i> IMS101	Trich101	β	7,75	34,14	5156	60,11	•	M	S	III

3.1.2. Cluster von vermutlich orthologen Genen (CLOG)

Um die genetische Diversität zu untersuchen, wurden Cluster von vermutlich orthologen Genen (*Cluster of Likely Ortholog Genes*) (CLOG) bestimmt, indem alle identifizierten Proteinsequenzen aller 16 cyanobakteriellen Stämme paarweise, „jede gegen jede“ verglichen wurden. Um für jede mögliche Kombination der Spezies A und B die orthologen Gene zu identifizieren, wurde eine Methode angewendet, die ähnlich der KEGG-Datenbank ist beziehungsweise des automatisierten Annotationsservers von KEGG [191]. Die Analyse beschränkte sich nur auf die Chromosomen der cyanobakteriellen Genome. Ihre Plasmide wurde nicht berücksichtigt.

Zwei Proteinsequenzen wurden als vermutliche Orthologe betrachtet, wenn beim reziproken Vergleich die bidirektionale Trefferrate (*Bidirectional Hit Rate*) (BHR) größer war als ein gegebener Schwellenwert. Dafür wurden zuerst alle Gene von Spezies A mit jedem Gen aus Spezies B verglichen und umgekehrt. Für diesen Vergleich wurde der BLASTp –Algorithmus benutzt. Alle Treffer, die einen Bit-Score kleiner als 50 Bits aufwiesen, wurden nicht weiter berücksichtigt. Anschließend wurde mithilfe des Bit-Scores eine BHR wie folgt berechnet:

$$BHR = \left(\frac{S_{a,b}}{S_b^{bestA}} \right) \times \left(\frac{S_{b,a}}{S_a^{bestB}} \right),$$

wobei $S_{a,b}$ der BLASTp-Score von a gegen b ist und S_b^{bestA} der beste Score von b gegen irgendein Gen in A ist (welches nicht a sein muss).

Der Wert der BHR ist für zwei Gene gleich Eins, wenn sie sich gegenseitig den besten Treffer in beide Richtungen liefern, andernfalls ist er niedriger. Überdies wurden auch für die Gene, die auf demselben Genom lokalisiert sind, die BHR errechnet. Um aber genomübergreifende orthologe Gene im weiteren Verlauf zu begünstigen, wurde für die Analyse die BHR von Genen, die auf demselben Genom liegen, auf einen Wert von 0,95 beschränkt, auch wenn der errechnete Wert höher liegen konnte. Alle Genpaare mit einer BHR größer oder gleich 0,95 wurden als mutmaßliche Orthologe klassifiziert.

In einem zweiten Schritt wurden die Genpaare in vorläufige Cluster gruppiert, indem alle Gene, die als vermutliche Orthologe identifiziert wurden, zusammengefasst wurden. Um Cluster zu vermeiden, in denen zwei Gene eine niedrige BHR aufweisen, aber trotzdem eine schwache Verbindung durch ein drittes Gen besitzen, wurden alle Gene in diesem vorläufigen Cluster mithilfe der UPGMA (*Unweighted Pair Group Method with Arithmetic mean*) und einer minimalen BHR von 0,75 erneut gruppiert [192]. Dies wurde erreicht, indem die engsten Gene, also die mit der höchsten BHR, zuerst gruppiert wurden und anschließend die BHRs zu allen anderen Einträgen des Clusters berechnet wurden:

$$BHR_{X,C} = \frac{|A| \times BHR_{A,C} + |B| \times BHR_{B,C}}{|A| + |B|},$$

wobei X die neue Einheit bestehend aus A und B ist, $|A|$ die Größe von A und $BHR_{X,C}$ die BHR von zwei Einträgen, bis alle Paare von Einheiten eine BHR größer oder gleich 0,75 besitzen. Bei diesem Verfahren wurde jedes Gen nur einem einzigen Cluster zugeordnet.

Der Ansatz der Analyse basiert auf schon früher publizierten Untersuchungen, die das Ziel hatten, CLOGs in mehrere Genomsequenzen zu identifizieren [193–200]. Jedoch wurden in der Analyse in dieser Arbeit eher strengere Kriterien festgesetzt, um die Aufnahme von falschen, nicht orthologen Genpaaren zu vermeiden (falsche Positive). Dies hat aber zur Folge, dass die Anzahl an echten potenziellen orthologen Genen unterschätzt wird. Diese Analyse wurde von Christan Beck durchgeführt.

3.1.3. Anreicherung der GO-Annotation

Um funktionale Unterschiede zwischen Kern- und einzigartigen CLOGs zu untersuchen, wurde jedem CLOG die jeweilige Funktionsannotation für jedes Gen in dem CLOG aus der GO-Datenbank (*Gene Ontology*) [201, 202] (Stand: Januar 2011) zugewiesen. Für die Analyse bezüglich welche GO-Einteilungen vermehrt auftreten, wurde die TopGO-Software verwendet [203], welche Teil des Bioconductor R-Paketes (<http://www.bioconductor.org>) ist. Hierfür wurde der Eltern-Kind-Algorithmus [204] mit Standardparametern und exaktem Fisher-Test gewählt. Für die Berechnung der p-Werte wurden nur CLOGs mit einer zugeordneten GO-Einteilung berücksichtigt. Diese Analyse wurde von Christan Beck durchgeführt.

3.1.4. Phylogenetische Analyse

Die phylogenetischen Bäume, die in dieser Arbeit gezeigt werden, wurden mit dem PHYLIP-Paket von Felsenstein [205] (*PHYLogeny Inference Package* Version 3.69) erstellt. Für den 16S RNA Vergleich (Abbildung 3.4 A), wurde ein Baum im Newick-Format unter Nutzung der *Ribosomal Database Project*-Webseite (<http://rdp.cme.msu.edu/>) [206] konstruiert. Anschließend wurde der Baum mit der DRAWGRAM-Funktion aus dem PHYLIP-Paket gezeichnet.

Für den Baum der gemeinsamen Cluster (Abbildung 3.4 B) wurde zuerst eine Ähnlichkeitsmatrix für alle Stämme erstellt, wobei die Ähnlichkeit zwischen zwei Stämmen durch die Anzahl der gemeinsamen Cluster geteilt durch die Gesamtanzahl der Cluster, in denen mindestens einer der beiden Stämme gelistet ist, definiert ist. Die Kern-Cluster, welche für alle Stämme gleich sind und die einzigartigen Clustern wurden für die Ähnlichkeitsmatrix nicht berücksichtigt. Letzteres diente auch dazu, eine Verzerrung der Ergebnisse zu minimieren. Aus der Ähnlichkeitsmatrix wurde anschließend eine Distanzmatrix berechnet, indem jeder Eintrag vom maximalen Eintrag subtrahiert wurde. Diese Distanzmatrix wurde abschließend in einen Baum mittels NEIGHBOR-Funktion aus dem PHYLIP-Paket umgewandelt und mit der DRAWGRAM-Funktion gezeichnet. Für beide Bäume wurde Aca11017 als Außengruppe verwendet. Diese Analyse wurde von Christan Beck durchgeführt.

3.1.5. Zuweisung von metabolischen Funktionen zu den CLOGs

Um den zuvor berechneten CLOGs eine Stoffwechselfunktion zuzuweisen, wurden die in den CLOGs vorkommenden Genen mit der KEGG-Datenbank (Stand: 19. Oktober 2010) [51] abgeglichen. CLOGs mit mindestens einem Gen, welches mit einer enzymatischen Funktion in der KEGG-Datenbank assoziiert ist, wurden mit der jeweiligen EC-Nummer gekennzeichnet. Für den Fall, dass ein CLOG mehrere Gene mit unterschiedlichen EC-Nummern beinhaltet, wurde der CLOG mit mehreren EC-Nummern versehen. Nicht berücksichtigt wurden dabei EC-Nummern, die nur eine unvollständige Variante einer anderen EC-Nummer sind (zum Beispiel 3.7.- von 3.7.4.21). Folglich muss die Gesamtanzahl der unterschiedlichen EC-Nummern nicht genau der Anzahl der metabolischen CLOGs entsprechen. Bei Enzym-Komplexen, die aus mehre-

ren Untereinheiten bestehen, welche durch mehrere einzelne Gene kodiert sind und die deshalb zu verschiedenen CLOGs zugeordnet sind, wurde jeder zu dem Enzym gehörige CLOG mit der EC-Nummer des entsprechenden Enzyms gekennzeichnet.

Im Kontext dieser Analyse ist darauf hinzuweisen, dass EC-Nummern nicht nur streng zu metabolischen Enzymen zugewiesen sind, sondern auch Enzyme umfassen, die eine allgemeine enzymatische Aktivität besitzen, wie zum Beispiel DNA- und RNA-Polymerasen oder Proteinkinasen.

3.2. Ergebnisse und Diskussion

3.2.1. Genomanalyse und Cluster von vermutlich orthologen Genen

Ausgangspunkt der Analyse waren die Genomsequenzen von 16 ausgewählten Cyanobakterien, die in der GenBank-Datenbank (<http://www.ncbi.nlm.nih.gov/genbank>) verfügbar sind. Die gewählten 16 Stämme sind nicht auf ein einziges Genus der Cyanobakterien beschränkt, sondern wurden in der Weise ausgewählt, dass sie die genomische und metabolische Diversität reflektieren, die unter dem Phylum der Cyanobakterien zu finden ist. Dementsprechend wurden acht marine und acht Süßwasserstämme selektiert. Zu den 16 ausgewählten Cyanobakterienstämmen gehören die Modellorganismen *Synechocystis* sp. PCC 6803, *Synechococcus elongatus* PCC 7942 und *Cyanotheca* ATCC 51142, mehrere stickstofffixierende (diazotrophe) Cyanobakterien sowie zwei thermophile Cyanobakterien, die ursprünglich aus der Umgebung von heißen Quellen isoliert wurden. Details zur Auswahl der Stämme sind in Kapitel 3.1.1 beschrieben und eine Zusammenfassung der Eigenschaften und Charakteristiken der einzelnen ausgewählten Stämme sind in Tabelle 3.1 dargestellt.

Der benutzte Algorithmus erkannte 21.238 unterschiedliche Cluster von vermutlich orthologen Genen (*Cluster of Likely Ortholog Genes*), verteilt über alle 16 Stämme. Abbildung 3.1 zeigt ein Histogramm der Anzahl zugewiesener Gene pro CLOG. Die Mehrzahl der Cluster (fast 60 %) enthält nur ein einziges Gen, während nur eine kleine Anzahl von Clustern mehr als 30 oder 40 Gene besitzt. Es fällt auf, dass CLOGs mit genau 16 Genen überrepräsentiert sind (siehe Abbildung 3.1, durch eine vertikale Linie markiert). Insgesamt unterscheidet sich die Verteilung leicht von den Ergebnissen, die in der COG-Datenbank (*Clusters of Orthologous Groups*) bereitgestellt

werden [200, 207]. In dieser werden nur zwei Cyanobakterienstämme aufgeführt (*Synechocystis* sp. PCC 6803 und *Nostoc* sp. PCC 7120). Betrachtet man diese beiden, neigen Cluster von orthologen Genen in der Datenbank dazu, auch aus mehreren Genen zu bestehen, die meist aus dem gleichen Stamm kommen.

Um einen Einblick in die Struktur der genomischen Diversität unter den Cyanobakterien zu erhalten, wurde jeder CLOG einem Cyanobakterienstamm zugewiesen, wenn ein oder mehrere Gene des CLOGs im jeweiligen cyanobakteriellen Genom zu finden waren: Abbildung 3.2 A zeigt ein Histogramm der Anzahl von CLOGs als Funktion der Anzahl der dazu assoziierten Stämme. Man kann zwischen Kern-Genen

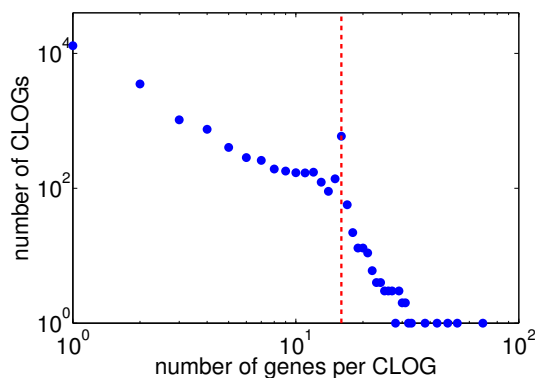


Abbildung 3.1.: Anzahl der Gene pro CLOG. Die Mehrzahl der CLOGs beinhaltet ein einziges Gen. CLOGs mit 16 Genen, hervorgehoben durch eine vertikale Linie, sind überrepräsentiert. Nur sehr wenige Cluster bestehen aus mehr als 16 Genen und nahezu kein Cluster beinhaltet mehr als 32 Gene (Abbildung aus Beck et al. (2012) [130]).

(660 CLOGs), welche allen 16 Stämmen zugeordnet sind, gemeinsamen Genen (6668 CLOGs), die in mehr als einem Stamm vorhanden sind, aber nicht in allen, und einzigartigen Genen (13.910 CLOGs), die keine vermutlich orthologen Gene in einer der anderen 15 Genomsequenzen besitzen, unterscheiden.

In Abbildung 3.2 B sind die Anzahl der CLOGs, die jeweils jedem Stamm zugeordnet wurden, dargestellt. Dabei sind die Anteile von Kern-, gemeinsamen und einzigartigen CLOGs eingezeichnet. Es ist zu erkennen, dass die Mehrzahl der orthologen Cluster nur mit einem einzelnen Genom assoziiert ist. Sie repräsentieren einzigartige Gene ohne vermutliche Orthologe in irgendeinem anderen der 15 Stämme. Zudem kann man feststellen, dass die Anzahl der CLOGs, welche zwei oder mehr Genome gemeinsam haben, dann schnell abfällt. Hierbei muss beachtet werden, dass die Anzahl der CLOGs in Abbildung 3.2 A logarithmisch skaliert ist. Nichtsdestotrotz ist eine wesentliche Anzahl von CLOGs (660 CLOGs) dem Kern-Genom zugewiesen. Die Anzahl der Kern-CLOGs stimmt ziemlich genau mit den Ergebnissen aus Mulkidjanian et al. (2006) [198] überein. Insbesondere wenn man *Synechocystis* als Referenz verwendet, findet man, dass fast alle Gene (> 90 %), welche in der Analyse einem Kern-CLOG zugewiesen wurden, ebenfalls im cyanobakteriellen Kern-Cluster von Mulkidjanian et al. (2006) [198] zu finden sind. Auch mit Ergebnissen aus mehreren früheren Studien über andere bakterielle Spezies sind die Resultate der Analyse qualitativ gut zu vereinbaren. Zum Beispiel haben Hogg et al. (2007) [195] eine ähnliche Verteilung für 12 sequenzierte Stämme von *Haemophilus influenzae* beobachtet. Untersucht man das Konzept des Pan-Genoms im größeren taxonomischen Maßstab, so berichten Lapierre und Gogarten (2009) [208] von einem gemeinsamen Kern-Genom, welches ca. 250 Gene aus mehr als 500 sequenzierten bakteriellen Genomen beinhaltet. In beiden Fällen wurde, genau wie die Ergebnisse in Abbildung 3.2 zeigen, eine U-förmige Verteilung beobachtet, so dass die einzigartigen und die Kern-Gene überrepräsentiert sind.

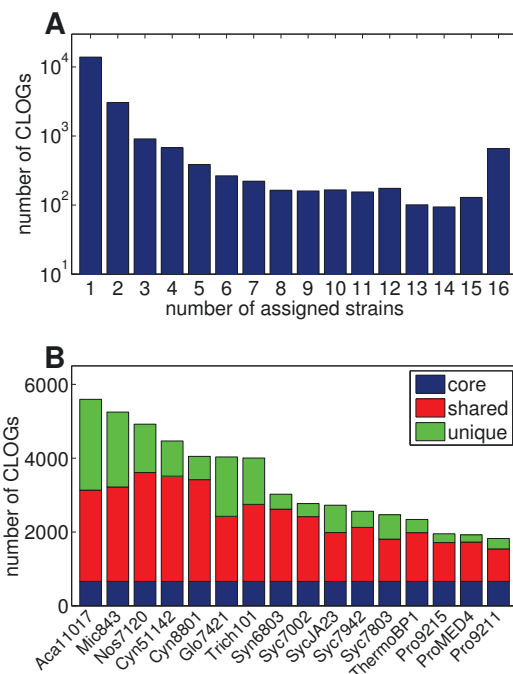


Abbildung 3.2.: Verteilung der CLOGs innerhalb der 16 untersuchten Cyanobakterienstämme. A) Ein Histogramm der Anzahl der CLOGs, die einer bestimmten Anzahl von Stämmen zugewiesen wurden. Zu erkennen ist die Anzahl der Kern-CLOGs (660, die bei allen 16 Stämmen vorhanden sind), der geteilten CLOGs (6668, bei 2-15 Stämmen) und der einzigartigen CLOGs (13910, nur in einem einzigen Stamm auftretend). B) Anzahl der CLOGs pro Cyanobakterienstamm. Hierbei sind die Anteile der Kern-, gemeinsamen und einzigartigen CLOGs farblich hervorgehoben (Abbildung aus Beck et al. (2012) [130]).

3.2.2. Das cyanobakterielle Kern- und Pan-Genom

Vergleichsanalysen auf Genomebene bieten die Möglichkeit, die beobachteten Fakten und Ergebnisse, die sich explizit für die untersuchten Organismen beziehungsweise Stämme ergeben, auf eine viel größere Anzahl von Organismen beziehungsweise Stämmen, die in der Vergleichsanalyse nicht berücksichtigt wurden, zu extrapolieren. Dies hat zur Folge, dass sich in den letzten Jahren die Pan-Genomanalyse als neuartiger Ansatz, um die Größe des Genrepertoires unterschiedlicher Spezies abzuschätzen, entwickelt hat [209]. Übereinstimmend haben mehrere neuere Studien festgestellt, dass für gewöhnlich die Anzahl der Gene, welche für eine bakterielle Spezies verfügbar sind, um mehrere Größenordnungen größer ist als die Anzahl von Genen im Genom eines einzelnen Organismus. Diese Erkenntnisse haben eine direkte Auswirkung auf die Genomsequenzierung von Organismen, da nun die Möglichkeit besteht, bei jeder Sequenzierung eine Vorhersage über die Anzahl der vermutlich neu identifizierten Gene eines neuen Organismus zu treffen. In Abbildung 3.3 ist eine Schätzung der Größe des cyanobakteriellen Kern- und des Pan-Genoms durch die analysierten 16 Cyanobakterien-Stämme dargestellt.

Das gesamte Pan-Genom aller 16 Stämme umfasst mehr als $2 \cdot 10^4$ orthologe Cluster und eine Erhöhung der Anzahl der analysierten Genome zeigt keine wesentliche Abflachung der Kurve (siehe Abbildung 3.3 B). Es ist zu erkennen, dass mit jedem neu aufgenommenen Genom das Pan-Genom um mehr als 500 neue orthologe Cluster erweitert wird. Somit kann erwartet werden, dass durch die Sequenzierung von weiteren cyanobakteriellen Stämmen eine große Anzahl von neuen, noch unbekannten Genen zu erwarten ist.

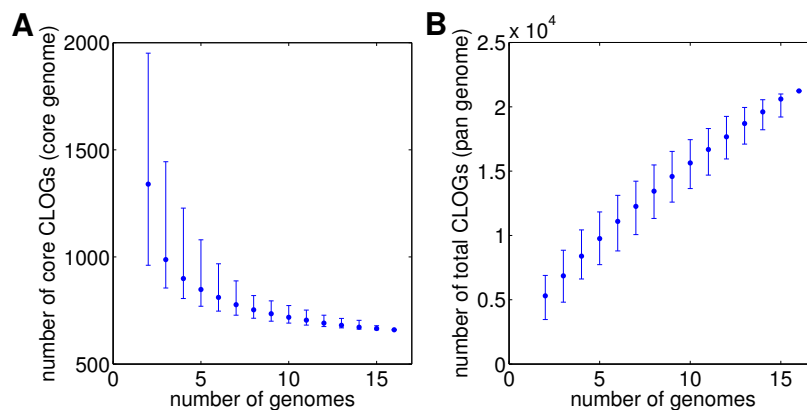


Abbildung 3.3.: Schätzung der Größe des cyanobakteriellen Kern-(A) und Pan-(B) Genoms. Um einen Einfluss der Stammreihenfolge zu verhindern, wurden die 16 Cyanobakterienstämme in zufälliger Reihenfolge hinzugefügt. Bei jedem Schritt wurde die Anzahl der Kern-CLOGs und der Pan-CLOGs (die Gesamtheit aller CLOGs) neu berechnet. Dieser Vorgang wurde über 1.000 Mal wiederholt. Gezeigt ist der Durchschnittswert dieser Wiederholungen. Die Fehlerbalken entsprechen den Quantilen von 0,1 und 0,9 bei 1.000 Wiederholungen (Abbildung aus Beck et al. (2012) [130]).

Nun ergeben sich natürlich aus den Abbildungen 3.2 und 3.3 zwei Fragen:

- Welchen Umfang hat das gesamte cyanobakterielle Pan-Genom?
- Worin liegt der funktionale und evolutionäre Unterschied zwischen den Kern-, gemeinsamen und einzigartigen Genen, wenn einer existiert?

Beide Fragen wurden auch schon in der Literatur behandelt, konnten jedoch bis heute nicht mit Sicherheit geklärt werden.

Bezüglich der Größe des bakteriellen Pan-Genoms weichen die derzeitigen Ergebnisse für verschiedene Arten stark voneinander ab. Hogg et al. (2007) [195], berichten über einen endlichen Umfang des Pan-Genoms für *Haemophilus influenzae*, extrapoliert aus 12 Genomsequenzen, während Ergebnisse für *Streptococcus agalactiae* auf ein unendliches Pan-Genom hindeuten [210].

In der Tat könnten diese Resultate die Folge von Unterschieden in der Evolutionsgeschichte und in den Lebensräumen beziehungsweise ökologischen Nischen der verschiedenen Arten sein. Anzumerken ist jedoch, dass erst kürzlich ein grundsätzlicher Einwand zur mathematischen Extrapolation von Kislyuk et al. [196] eingebracht wurde. Diese argumentieren, dass eine Extrapolation von seltenen Ereignissen, wie das Auftreten seltener Gene und Genome höchstwahrscheinlich eine Vielzahl von falschen Einschätzungen liefert. Deshalb wurde in dieser Arbeit keine Schätzung beziehungsweise Aussage für die Größe des gesamten cyanobakteriellen Pan-Genoms getroffen. Dennoch konnten mehrere neue wichtige Erkenntnisse gewonnen werden:

- Es existiert ein Kern-Genom, welches von diesen 16 Cyanobakterienstämmen geteilt wird.
- Durch Extrapolation ist es derzeit nicht möglich, die asymptotische Größe des Kerns-Genoms aller Cyanobakterienstämme abzuschätzen .
- Weiterhin gibt es keinen Hinweis, dass das cyanobakterielle Pan-Genom endlich beziehungsweise geschlossen ist.

Daher liefern die Ergebnisse aus Abbildung 3.3 einen großen Anreiz für weitere Genomsequenzierungen, selbst von eng verwandten Stämmen.

Ein weiteres Problem betrifft die möglichen funktionellen und evolutionären Unterschiede zwischen den Kern-, gemeinsamen und einzigartigen Genen. Alle neueren Studien haben gemeinsam, dass die Anzahl der einzigartigen Gene und diejenigen, die nur eine kleine Anzahl der Genome gemeinsam haben, einen recht großen Anteil des gesamten Genrepertoires darstellen [211]. Es existiert eine Vielzahl von Hypothesen bezüglich des Ursprungs einer solchen Verteilung. Zum Beispiel wird oft angenommen, dass Kern-Gene sich überwiegend den Housekeeping-Genen zuordnen lassen [211]. Andererseits können einzigartige Gene als charakteristisch für verschiedene Umgebungen angenommen werden und die Folge eines intensiven horizontalen Gentransfers sein [212, 213].

Diese Behauptungen wurden durch den Vergleich der Annotation, die man aus der *Gene Ontology* (GO)-Datenbank [201] erhält, getestet. Eine Analyse der GO-Annotation von Kern-CLOGs offenbarte eine signifikante Anreicherung von Genen mit Bezug auf die Kategorie „Translation“ (p-Wert $< 1 \cdot 10^{-30}$), „DNA-Reparatur“ (p-Wert $< 1 \cdot 10^{-4}$), „Genexpression“ (p-Wert $< 1,8 \cdot 10^{-7}$), „RNA-Prozessierung und Modifizierung“ (p-Wert $< 1 \cdot 10^{-5}$), diverse Transportprozesse (p-Werte $< 1 \cdot 10^{-4}$) sowie mehreren metabolischen und biosynthetischen Prozessen (p-Werte $< 1 \cdot 10^{-5}$). Gene, die in den einzigartigen CLOGs zu finden sind, besitzen meist die Annotation von „Abwehrreaktion“ (p-Wert $< 1,6 \cdot 10^{-5}$) oder „DNA-Integration“ (p-Wert $< 8,2 \cdot 10^{-5}$) und sind insbesondere häufig in der Annotation von regulatorischen Prozessen zu finden. Demzufolge implizieren Kern-Gene eine Erhaltung der Hauptfunktionen der Zelle wie Stoffwechsel und die Genexpression, während die regulatorischen Eigenschaften und Wechselwirkungen eher spezifisch für die Anpassung an diverse Umgebungen sind (einzigartige Gene).

Es muss darauf hingewiesen werden, dass es eine mögliche Verzerrung gibt, da eine deutliche Anreicherung von GO-annotierten Genen in den Kern-CLOGs ($p < 9 \cdot 10^{-210}$ nach Fisher-Test) existiert, während Gene, die in den einzigartigen CLOGs zu finden sind, eher keine GO-Annotation besitzen. Diese Unausgewogenheit kann durch die Tatsache erklärt werden, dass die GO-Annotationen hauptsächlich auf BLAST-Suchen bei anderen Organismen basieren und einzigartige Gene erwartungsgemäß weniger positive BLAST-Treffer besitzen. Ferner könnte sich auch ein großer Anteil der einzigartigen Gene aus Falschannotation beziehungsweise Annotations-Fehlern ergeben haben oder es sind nicht funktionale Gene, die aufgrund des fortlaufenden Prozesses der Genomreduktion auftreten [211].

3.2.3. Phylogenetische Übereinstimmungen zwischen den 16 Cyanobakterienstämmen

Um die cyanobakteriellen Spezies zu vergleichen und deren Beziehungen untereinander zu entschlüsseln, sind in Abbildung 3.4 zwei phylogenetische Bäume gezeigt. Ein konventioneller Stammbaum, basierend auf der 16S rRNA, berechnet und gezeichnet mit dem PHYLIP-Paket von Felsenstein [205] (Abbildung 3.4 A) und dazu im Vergleich ein phylogenetischer Baum, berechnet aus den Ergebnissen des vorher beschriebenen Genvergleichs.

Für die Schätzung der Ähnlichkeit von den Spezies untereinander auf Grundlage der CLOGs stehen mehrere Optionen zur Verfügung. Bei dieser Analyse wurde ein einfaches Maß, basierend auf der Anzahl von CLOGs, die zwei Stämme gemeinsam haben, geteilt durch die Gesamtzahl der CLOGs, die beiden Stämmen zusammen zugeordnet sind, zugrunde gelegt. Der daraus errechnete Distanzbaum ist in Abbildung 3.4 B gezeigt. Beide Bäume weisen ein hohes Maß an Ähnlichkeit auf, mit nur wenigen topologischen Unterschieden.

In beiden Fällen bilden die *Prochlorococcus*-Stämme die am engsten verwandte Gruppe. Hierbei ist anzumerken, dass phylogenetische Bäume einzelner Genfamilien, in denen ein höherer Grad der phylogenetischen Verschiedenheit erwartet wird [214],

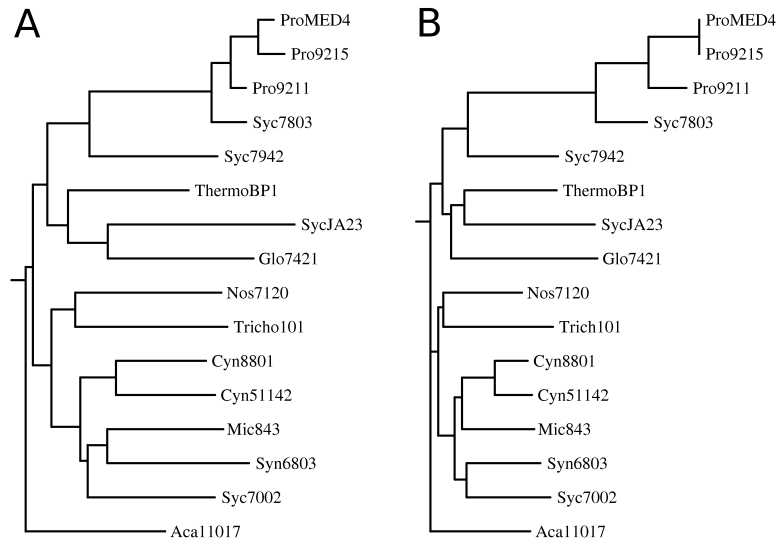


Abbildung 3.4.: Phylogenetische Analyse der 16 untersuchten Cyanobakterienstämme. A) Phylogenetischer Baum basierend auf einem Vergleich auf Ebene der 16S rRNA. B) Phylogenetischer Baum basierend auf der Anzahl der gemeinsamen CLOGs, die von zwei Stämmen geteilt werden (Abbildung aus Beck et al. (2012) [130]).

bei dieser Analyse nicht betrachtet werden. Ebenso ist jede Schätzung der Distanz auf Basis der gemeinsamen CLOGs durch die unterschiedliche Genomgröße voraussichtlich fehlerbehaftet, was wiederum die evolutionäre Distanz, bestimmt durch 16S-rRNA-Analyse, widerspiegelt.

In Tabelle 3.2 ist ein paarweiser Vergleich der gemeinsamen CLOGs zwischen allen 16 Cyanobakterienstämmen zu sehen. Auch diese Tabelle bestätigt erneut die enge Verbindung der drei *Prochlorococcus*-Stämme und Syc7803 in Bezug auf die gemeinsamen Gene.

Tabelle 3.2.: Tabelle der geteilten CLOGs. Jeder Eintrag enthält die Anzahl der CLOGs, die von zwei Stämmen geteilt werden. Die unterste Zeile zeigt jeweils die gesamte Anzahl der CLOGs eines Stammes (Tabelle aus Beck et al. (2012) [130]).

	Cyn51142	Cyn8801	Glo7421	Mic843	Nos7120	ProMED4	Pro9211	Pro9215	SycJA23	Syc7002	Syc7803	Syc7942	Syn6803	ThermoBPI	Trich101
Aca11017	2023	1962	1574	1841	2167	1016	1018	997	1454	1736	1228	1602	1726	1572	1765
5592															
	Cyn51142	2721	1561	2285	2379	988	1008	979	1481	1903	1238	1625	2049	1557	1980
4465															
	Cyn8801	1537	2301	2265	983	1004	970	1474	1871	1208	1622	2043	1537	1916	
	4048														
	Glo7421	1546	1741	868	892	853	1333	1334	1037	1278	1385	1269	1378		
	4082														
	Mic843	2216	988	1000	963	1450	1788	1186	1626	1951	1517	1863			
	5247														
	Nos7120	989	1000	974	1596	1911	1228	1671	1906	1614	2079				
	4920														
	ProMED4	1334	1616	904	973	1277	1022	972	924	974					
	1922														
	Pro9211	1309	908	983	1340	1040	986	944	992						
	1825														
	Pro9215	888	962	1285	998	961	910	964	964						
	1951														
	SycJA23	1380	1055	1345	1379	1364	1418								
	2725														
	Syc7002	1169	1547	1812	1459	1635									
	2771														
	Syc7803	1265	1169	1109	1182										
	2467														
	Syc7942	1581	1441	1491											
	2561														
	Syn6803	1490	1644												
	3024														
	ThermoBPI	1450													
	2340														
	Trich101	1404													
	4004														

3.2.4. Das metabolische Netzwerk der Cyanobakterien ist stark konserviert

Natürlich liegt neben der Pan-Genomanalyse auch ein großes Interesse auf der Struktur und der Diversität des cyanobakteriellen Stoffwechsels. Um die CLOGs zu identifizieren, die mit einer metabolischen Funktion assoziiert sind, wurden die *Enzyme Commission* (EC)-Nummern, die für jedes einzelne Gen in der KEGG-Datenbank verfügbar waren, benutzt. Dabei wurde ein CLOG als metabolisch klassifiziert, wenn zu der jeweiligen Gruppe von orthologen Genen eine oder mehrere EC-Nummern mit einer spezifischen enzymatischen Aktivität zugeordnet werden konnten. Hierbei ist anzumerken, dass die hierarchische Systematik der EC-Nummern bei der Zuordnung sowohl breitere enzymatische Kategorien als auch eine begrenzte Anzahl von nicht metabolischen Enzymen umfasst. Bei der Anwendung dieser Kriterien auf die CLOGs aus dem Kapitel 3.2.1 ist auszumachen, dass von diesen 21.238 CLOGs, 1.851 als metabolische CLOGs betrachtet werden können. Zudem kann man feststellen, dass aufgrund von bifunktionalen Enzymen oder widersprüchlicher oder fehlerhafter Annotation, den CLOGs auch mehrere metabolische Funktionen zugeordnet werden können. Bei der Analyse in dieser Arbeit war dies bei nur 66 CLOGs (von den 1.851 metabolischen CLOGs) der Fall, bei denen einem CLOG mehr als eine EC-Nummer zugeordnet wurde. Somit ist zu erkennen, dass eine widersprüchliche oder fehlerhafte Annotation die Ergebnisse der Analyse, auch ohne eine vorhergehende Filterung oder manuelle Kuration, nicht wesentlich einschränken. Insgesamt wurden 759 verschiedene EC-Nummern den Clustern zugewiesen.

Abbildung 3.5 zeigt die Verteilung der metabolischen CLOGs über die einzigartigen, gemeinsamen und Kern-Gene der einzelnen cyanobakteriellen Genome. Die CLOGs mit einer metabolischen Funktion sind stark überrepräsentiert in den CLOGs, die alle 16 cyanobakteriellen Stämme gemeinsam haben. Rund 55 % aller Kern-CLOGs sind mit einer metabolischen Funktion assoziiert. Offenbar stellt der zelluläre Stoffwechsel, hier definiert als Gene mit einer metabolischen Funktion, einen großen Anteil des Kern-Genoms dar.

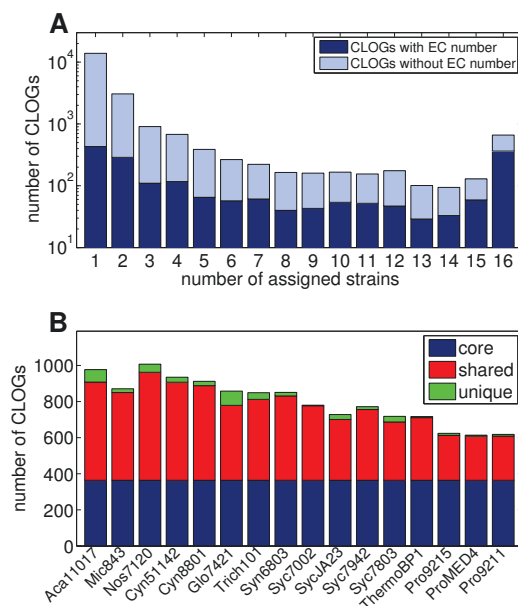


Abbildung 3.5.: Verteilung von metabolischen CLOGs innerhalb der cyanobakteriellen Genome. A) Ein Histogramm der Anzahl der CLOGs, die einer bestimmten Anzahl von Stämmen zugewiesen wurden (logarithmisch skaliert). Der Anteil an CLOGs mit einer metabolischen Funktion ist hierbei farblich hervorgehoben. Unter den Kern-CLOGs ist dieser Anteil besonders hoch. B) Ein Säulendiagramm der Anzahl der CLOGs mit einer metabolischen Funktion für jeden einzelnen Stamm mit Unterscheidung der CLOG-Kategorie (Abbildung aus Beck et al. (2012) [130]).

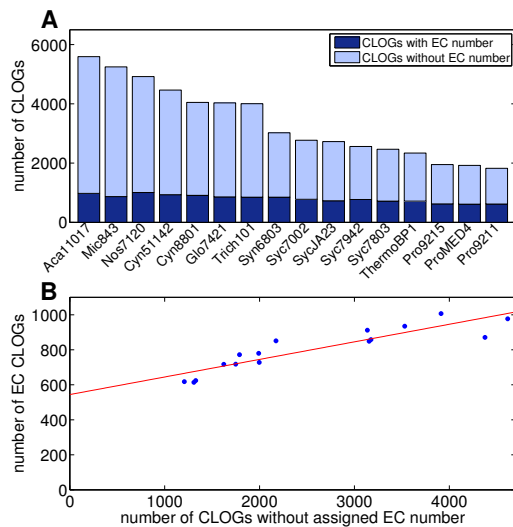


Abbildung 3.6.: Vergleich von metabolischen und nicht metabolischen CLOGs. A) Ein Säulendiagramm der Anzahl der CLOGs pro Stamm, die mit einer EC-Nummer assoziiert sind beziehungsweise nicht assoziiert sind. B) Anzahl der CLOGs mit einer EC-Nummer im Vergleich zu CLOGs ohne EC-Nummer. Jeder Punkt entspricht einem der 16 cyanobakteriellen Genome. Die rote Linie entspricht einer Regressionsgeraden (Methode der kleinsten Quadrate) um die Tendenz der Daten hervorzuheben. Es herrscht eine schwache Korrelation zwischen der Größe (der Gesamtanzahl der CLOGs) und der Anzahl der metabolischen CLOGs eines Stamms (Abbildung aus Beck et al. (2012) [130]).

In Abbildung 3.6 sind die prozentualen Werte der CLOGs, denen eine enzymatische Funktion zugewiesen wurde, von allen 16 Stämmen gezeigt. Ausgehend von etwa 500 metabolischen CLOGs des Kern-Genoms steigt die Anzahl der enzymatischen CLOGs linear mit der Anzahl der gesamten CLOGs, die jedem einzelnen Stamm zugeordnet werden, an. Allerdings ist die Korrelation zwischen der Anzahl der enzymatischen CLOGs und der Gesamtanzahl der CLOGs eher schwach und stark beeinflusst durch die relativ hohe Anzahl von enzymatischen Kern-CLOGs.

Eine weitere Analyse, ob bestimmte Stoffwechselwege oder Enzymklassen spezifisch verteilt sind, ergab keinen offensichtlichen Unterschied zwischen einzigartigen und Kern-Genen. Das heißt, in keiner der beiden Klassen ist eine besondere Enzym-Kategorie oder ein Stoffwechselweg stark überrepräsentiert. Trotzdem konnte eine gewisse Anzahl an Kern-Stoffwechselwegen identifiziert werden, die in allen 16 Cyanobakterienstämmen vorhanden sind. Unter den hochkonservierten Stoffwechsel-

wegen sind der Calvin-Benson-Zyklus, der oxidative Pentosephosphatweg, die Synthesewege von Nukleotiden und Aminosäuren zu finden. Auf letztere bezogen sind jedoch eine Reihe von Phosphatasen und Transaminasen in mehreren Cyanobakterienstämmen nicht annotiert.

3.2.5. Die Diversität des cyanobakteriellen Stoffwechsels

Schaut man sich bestimmte Enzyme beziehungsweise das metabolische Potenzial, welches unter diesen 16 Cyanobakterienstämmen verteilt ist, genau an, erhält man ein vielschichtiges und facettenreiches Bild. Zuerst wurde für diese Analyse die Anzahl der CLOGs auf gemeinsame CLOGs mit einer EC-Nummer beschränkt. Von den insgesamt 759 verschiedenen EC-Nummern aller Cluster, sind eine Teilmenge von 378 EC-Nummern mehr als einem, aber weniger als 16 Stämmen zugeordnet (gemeinsame CLOGs).

Abbildung 3.7 bietet eine geclusterte Heatmap der Zuordnung zu den einzelnen Stämmen dieser 378 gemeinsamen EC-Nummern. Hierfür wurden die EC-Nummern

unter Verwendung der MATLAB-Funktion *Clustergram* (mit Hamming-Distanz als Abstandsmaß) geclustert. Bei genauerer Betrachtung kann man zwischen vier Kategorien unterscheiden:

- Cluster A: gemeinsame EC-Nummern, die überwiegend in den *Prochlorococcus*-Stämmen mit Pro9215, ProMED4, Pro9211 und Syc7803 annotiert sind
- Cluster B: gemeinsame EC-Nummern, die nur für wenige Stämme annotiert sind
- Cluster C: gemeinsame EC-Nummern, die für eine große Anzahl an Stämmen annotiert sind
- Cluster D: gemeinsame EC-Nummern, die für fast alle Stämme annotiert sind, mit Ausnahme der drei *Prochlorococcus*-Stämmen und Syc7803

Es ist festzustellen, dass Abbildung 3.7 erneut die Ähnlichkeit zwischen den drei *Prochlorococcus*-Stämmen und Syc7803 hervorhebt, welche bereits in Abbildung 3.4 und Tabelle 3.2 zu erkennen ist.

Beim genaueren Betrachten der einzelnen Zuordnungen werden die Unterschiede zwischen diesen vier Gruppen noch ersichtlicher. Es wird deutlich, dass die beiden Enzyme Malat:Chinon-Oxidoreduktase (EC 1.1.5.4) (MQO), die im TCA-Zyklus involviert ist und die Cystathionin- γ -Synthase (EC 2.5.1.48), welche einen essentiellen Schritt bei der Synthese der Aminosäure L-Methionin katalysiert, exklusiv mit den vier Stämmen ProMED4, Pro9211, Pro9215 und Syc7803 assoziiert sind. Hingegen ist der Syntheseweg von L-Methionin bei den meisten anderen Stämmen noch ungeklärt. Es existieren keine EC-Nummern, die ausschließlich mit den drei *Prochlorococcus*-Stämmen

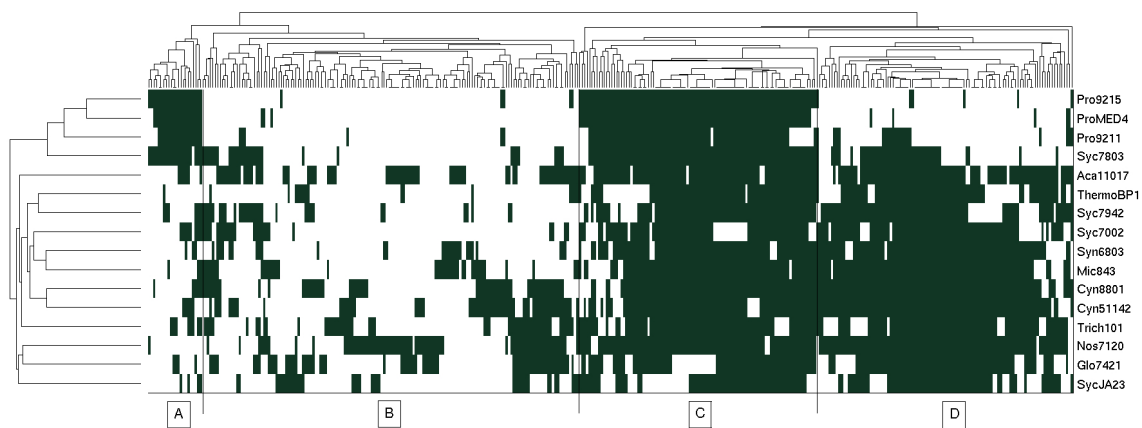


Abbildung 3.7.: Diversität des cyanobakteriellen Stoffwechsels. Gezeigt ist eine geclusterte Heatmap, die die Zuweisung der 378 gemeinsamen EC-Nummern zu den einzelnen Stämmen zeigt. Das Clusterschema hat eine hohe Übereinstimmung mit den Ergebnissen aus Abbildung 3.4. Mit Bezug auf die EC-Nummern kann durch einfaches Betrachten zwischen vier Kategorien unterschieden werden: Cluster A: EC-Nummern, die überwiegend bei den drei *Prochlorococcus*-Stämmen und Syc7803 auftreten. Cluster B: EC-Nummern, die nur bei wenigen Stämmen annotiert sind. Cluster C: EC-Nummern, die für eine große Anzahl an Stämmen annotiert sind. Cluster D: EC-Nummern, die bei fast allen Stämmen auftreten, aber wenig unter den *Prochlorococcus*-Stämmen und Syc7803 verbreitet sind (Abbildung aus Beck et al. (2012) [130]).

assoziiert sind. Umgekehrt enthält Gruppe D eine Reihe von 11 EC-Nummern, welche zu allen Stämmen außer den drei *Prochlorococcus*-Stämmen ProMED4, Pro9211, Pro9215 assoziiert sind. Dazu gehören unter anderem die Nitrit- und die Nitrate-Reduktase (EC 1.7.7.1 und EC 1.7.7.2). Dieses deckt sich mit der weit verbreiteten Annahme, dass Nitrat für die *Prochlorococcus* nicht verwendbar ist, da keiner der sequenzierten Laborstämme die entsprechenden Gene für eine Nitratverwertung besitzt [215]. Durch die hohe Verfügbarkeit von metagenomischen Sequenzdaten wurde diese Ansicht allerdings in letzter Zeit widerlegt, da herausgefunden wurde, dass doch mehrere andere *Prochlorococcus*-Ökotypen ein Gen für die Nitrat-Reduktase besitzen [216].

Unter den EC-Nummern, die nicht für Syc7803 und die drei *Prochlorococcus*-Stämmen annotiert sind, treten viele Enzyme des Zentralstoffwechsels auf, wie z.B. das Malatenzym (EC 1.1.1.38) (ME), die Glycerol-3-Phosphat-Dehydrogenase (EC 1.1.1.94) und die Fructose-1,6-Bisphosphatase (EC 3.1.3.11) (FBP).

3.2.6. Der Stoffwechsel von unterschiedlichen Speicherstoffen bei Cyanobakterien

Cyanobakterien zählen bekanntermaßen zu den phototrophen Organismen und verlassen sich demzufolge auf intrazelluläre Verbindungen, die als Speicherstoffe dienen und die Aufrechterhaltung von zellulären Prozessen während der Abwesenheit von Licht und vor allem in der Nacht gewährleisten.

Die häufigste Speicherstoffverbindung bei Cyanobakterien ist Glykogen, welches als verzweigtes Polymer aus Glucose-6-Phosphat synthetisiert wird. Es wird angenommen, dass Glykogen im Laufe des Tages, wenn genügend Licht vorhanden ist, aufgebaut wird und bei Abwesenheit von Licht wieder mobilisiert wird. Zudem ist der Glykogenstoffwechsel auch für das Wachstum unter Stickstoffmangelbedingungen von Bedeutung. Alle 16 Stämme, die in dieser Arbeit untersucht wurden, besitzen die notwendigen Enzyme zur Synthese und Mobilisierung von Glykogen. Insbesondere existiert ein CLOG, welcher für die ADP-Glucose-Pyrophosphorylase (EC 2.7.7.27) (AGP) annotiert ist, der auch zum Kern-Genom gehört und dementsprechend auch mit allen 16 analysierten Stämmen assoziiert ist. Genauso sind die Enzyme, wie die Glykogen-Synthase (EC 2.4.1.21) (GS) und das 1,4- α -Glucan-verzweigende Enzym (*Glycogen Branching Enzyme*) (EC 2.4.1.18) (GBE) für alle 16 Cyanobakterien-Stämme annotiert. Jedoch sind für diese beiden Stoffwechselschritte die jeweiligen Enzyme in verschiedene CLOGs aufgeteilt und aus diesem Grund nicht unbedingt als Orthologe für alle Stämme definiert. Derselbe Sachverhalt gilt auch für die Glykogen-Phosphorylase (EC 2.4.1.1) (GP), welche für die spätere Mobilisierung von Glykogen verantwortlich ist. Sie ist für alle 16 Stämme annotiert, allerdings nicht in einem einzigen CLOG. Ein Überblick über die entsprechenden Gene ist in Tabelle 3.3 zu finden.

Besonders zu erwähnen ist, dass in allen Fällen, in denen ein metabolisches Enzym mit mehr als einem CLOG assoziiert ist, normalerweise ein „primärer“ CLOG existiert, der mit fast allen Stämme verbunden ist und ein oder eine kleine Anzahl von „sekundären“ CLOGs, die mit demselben Enzym assoziiert sind. Bei einer genaueren

Betrachtung dieser Analyse wird deutlich, dass dies nicht ein Artefakt des Clustering-Algorithmus ist, sondern auch durch paarweise Vergleiche der jeweiligen Sequenzen unterstützt wird.

Im Vergleich zu Glykogen sind andere Speicherstoffverbindungen unter den Cyanobakterien weniger verbreitet. Nichtsdestotrotz sind die Enzyme für die Cyanophycin-Synthese und -Mobilisierung bei vielen Stämmen annotiert. Cyanophycin ist ein Polymer, welches zu gleichen Teilen aus den Aminosäuren L-Aspartat und L-Arginin gebildet wird. Somit dient es in mehreren Cyanobakterien sowohl als Stickstoff- als auch als Kohlenstoffquelle. Auch in diesem Fall ist das Enzym für die Cyanophycin-Synthetase (EC 6.3.2.29/30) (CphA) (siehe Knoop et al. [93]), zu mehreren CLOGs assoziiert, während das Enzym für die Cyanophycin-Mobilisierung, die Cyanophycina-se (EC 3.4.15.6) (CphB), nur für einen einzigen CLOG annotiert ist. Beide Enzyme sind in der Regel zusammen für einen Stamm annotiert. Das heißt, dass kein Stamm ein Enzym lediglich für die Synthese oder nur für die Mobilisierung besitzt. Jedoch existiert mit Syc7803 eine einzige Ausnahme. Für diesen Fall ist es aber sehr wahrscheinlich, dass maßgeblich eine falsche Annotation dafür verantwortlich ist, da es sich hierbei um einen einzelnen separaten CLOG handelt. Das zugehörige Gen ist als vermeintliche Cyanophycin-Synthetase in CyanoBase [114] annotiert, jedoch ist die Ähnlichkeit zu anderen bekannten Genen der Cyanophycin-Synthetase sehr niedrig.

Weniger verbreitet als Cyanophycin ist die Nutzung von Poly- β -Hydroxybutyrat (PHB) als Speicherstoff für Kohlenstoff. PHB ist ein nicht toxischer, biologisch abbaubarer Polyester mit großer Bedeutung für biotechnologische Anwendungen. Die Produktion von PHB durch gentechnisch veränderte Cyanobakterien wurde erst vor

Tabelle 3.3.: Zusammenfassung der Gene der 16 Cyanobakterienstämme, die in der Synthese und Mobilisierung von den drei verschiedenen Speicherstoffen Glykogen, Cyanophycin und Poly- β -Hydroxybutyrat involviert sind. Schwarze Punkte repräsentieren die Anzahl der Gene, die mit einem CLOG, dem eine spezifische enzymatische Aktivität zugewiesen wurde, assoziiert sind. Jede Spalte entspricht einem unterschiedlichen CLOG, jedoch können mehreren CLOGs die gleiche enzymatische Aktivität zugewiesen sein (Tabelle aus Beck et al. (2012) [130]).

	Glykogen				Cyanophycin		Poly- β -Hydroxybutyrat			
	AGP	GS	GBE	GP	CphA	CphB	PhaA	PhaB	PhaC	PhaE
Aca11017	••	• •	• - -	• - - -	- - -	-	•	-	-	-
Cyn51142	••	• • •	• • •	•• - - -	• - -	•	-	-	-	-
Cyn8801	•	• • •	• - •	•• - - -	• - -	•	-	-	-	-
Glo7421	•	• -	• • -	•• • - •	• - -	•	-	-	-	-
Mic843	•	• • •	• - -	• - - -	• - -	•	•	•	•	•
Nos7120	•	• • •	• • -	- • • -	• • -	•	-	-	-	-
ProMed4	•	• -	• - -	• - - -	- - -	-	-	-	-	-
Pro9211	•	• -	• - -	• - - -	- - -	-	-	-	-	-
Pro9215	•	• -	• - -	• - - -	- - -	-	-	-	-	-
SycJA23	•	• • •	• - -	- •• - -	• • -	•	-	-	-	-
Syc7002	•	• • •	• - -	• • - -	• - -	•	-	-	-	-
Syc7803	•	• -	• - -	• - - -	- - -	-	-	-	-	-
Syc7942	•	• -	• - -	• - - -	- - -	(•)	-	-	-	-
Syn6803	•	• • •	• - -	•• - - -	• - -	•	•	•	•	•
ThermoBP1	•	• -	• - -	•• - • -	• - -	•	-	-	-	-
Trich101	•	• -	• - -	•• - - -	• - -	•	-	-	-	-

kurzem in der Literatur diskutiert [217]. Von den 16 Cyanobakterienstämmen, die in dieser Analyse untersucht wurden, sind nur für *Synechocystis* und Mic843 die entsprechenden Enzyme zur Synthese von PHB annotiert. Es ist zu beachten, dass obwohl der Stamm Aca11017 ebenfalls ein Gen besitzt, welches mit dem CLOG, der für die Polyhydroxyalkanoat-spezifische β -Ketothiolase/Acetyl-CoA Acetyltransferase (EC 2.3.1.9) (PhaA) annotiert ist, assoziiert ist, die anderen verbleibenden Schritte für die PHB-Synthese fehlen. Der Grund hierfür könnte sein, dass das jeweilige Gen in Aca11017 eine eng verwandte Variante des PhaA-Gens ist, welches jedoch nicht spezifisch für die PHB-Synthese ist.

3.2.7. Die Diversität des Zentralstoffwechsels unter den Cyanobakterien

Neben den Speicherstoffen sind auch eine Reihe von anderen Enzymen des zentralen Stoffwechsels in Cyanobakterien von allgemeinem wissenschaftlichen Interesse. In Tabelle 3.4 und 3.5 sind mehrere wichtige Enzyme des zentralen Kohlenstoffmetabolismus der 16 untersuchten Cyanobakterienstämmen zusammengefasst. Im Gegensatz zu Tabelle 3.3 wurde bei diesen Tabellen nicht zwischen den einzelnen CLOGs, welche die gleiche enzymatische Funktion besitzen, unterschieden. Ähnlich zum Kapitel 3.2.6 existiert in der Regel für jedes Enzym ein primärer CLOG und eine kleine Menge an sekundären CLOGs.

Bei einer detaillierteren Analyse der jeweiligen metabolischen Funktion ist festzustellen, dass zunächst alle wichtigen Enzyme des Calvin-Benson-Zyklus, welcher die Fixierung von Kohlenstoffdioxid (CO_2) ermöglicht, in allen 16 Stämmen annotiert ist (siehe Tabelle 3.4). Ebenso sind für alle Enzyme des oxidativen Pentosephosphatweges (OPP) CLOGs für jeden der 16 Stämme assoziiert (siehe Tabelle 3.5). Für andere wichtige Stoffwechselwege ergibt sich jedoch ein differenzierteres Bild. Hier ist die Fructose-1,6-Bisphosphatase (EC 3.1.3.11) (FBP) zu nennen, die nicht für alle Stämme annotiert ist und in allen α -Cyanobakterien, einschließlich der *Prochlorococcus*-Stämme, fehlt. Jedoch zeigten schon die ersten Ergebnissen aus der Rekonstruktion des metabolischen Netzwerkes von *Synechocystis* (Knoop et al. [93, 94]) und deren Analyse mithilfe der Flussbilanzanalyse (*Flux Balance Analysis*), dass dieses Enzym für die Synthese von Biomasse nicht benötigt wird (siehe Kapitel 2.2).

Zum Teil kann auch dessen Funktion durch die bifunktionale Fructose-1,6-Bisphosphatase/ Sedoheptulose-1,7-Bisphosphatase (EC 3.1.3.37) (SBP) ersetzt werden, welche für alle 16 Stämme annotiert ist. Desgleichen ist die Phosphofructokinase (EC 2.7.1.11) (PFK) nicht für alle Stämme annotiert, sondern fehlt bei mehreren Stämmen, erneut vor allem bei *Prochlorococcus*. Zu erwähnen sei hier, dass die PFK essentiell für die Funktion der Glykolyse ist und bei einem Fehlen dieses Enzyms, die Verwendung von Glykogen als Kohlenstoff- und Energiequelle ausschließlich über den OPP möglich ist. Andere Enzyme der Glykolyse wie die Fructose-1,6-Bisphosphat-Aldolase (EC 4.1.213) (FBPA), Glyceraldehyd-3-Phosphat-Dehydrogenase (EC 1.2.1.12/59) (GAPDH), Phosphoglucomutase (EC 5.4.2.1) (PGM), und Pyruvatkinase (EC 2.7.1.40) (PYK) sind für alle 16 Stämme vorhanden (siehe Tabelle 3.4). Eine graphische Dar-

stellung des zentralen Stoffwechsels mit allen annotierten Stoffwechselwegen ist in Abbildung 3.8 zu sehen.

Im Unterschied zum Calvin-Benson-Zyklus ist der Pyruvatstoffwechsel, dessen Enzyme in Tabelle 3.5 zusammengefasst wurden, eher lückenhaft konserviert. Während CLOGs, welche mit der PEP-Carboxylase (EC 4.1.1.31) (PEPC) annotiert sind, bei allen 16 Stämmen vorhanden sind, ist hingegen die Rückreaktion über die PEP-Carboxykinase (EC 4.1.1.49) (PEPK) eher selten und nur für drei Stämme annotiert. Die PEPC katalysiert die Umwandlung von anaplerotischen PEP zu Oxalacetat und anorganischem Phosphat (P_i) und ist essentiell für die Regeneration des TCA-Zyklus.

Die Enzyme, die die Umwandlung von Oxalacetat zusammen mit Acetyl-CoA zu 2-Oxoglutarat katalysieren, sind bei allen 16 Cyanobakterienstämmen zu finden. Dieser Sachverhalt ist nicht sonderlich verwunderlich, da 2-Oxoglutarat ein Ausgangsstoff für die Synthese mehrerer Aminosäuren und Nukleotide ist und zudem als ein wichtiger Sensor für den Stickstoffhaushalt bei Cyanobakterien fungiert [218].

Jedoch sind die weiterführenden Schritte innerhalb des TCA-Zyklus hochgradig fragmentiert (siehe Tabelle 3.5 als Überblick). Es existieren keine CLOGs, die mit Genen des 2-Oxoglutarat-Dehydrogenase (OGDH)-Komplexes annotiert sind, welcher die Umwandlung von 2-Oxoglutarat zu Succinyl-CoA katalysiert. Das Fehlen eines OGDH stimmt mit dem schon seit Jahren bekannten Sachverhalt überein, dass Cyanobakterien keinen herkömmlichen TCA-Zyklus besitzen. Im Allgemeinen wird dennoch angenommen, dass der TCA-Zyklus bei Cyanobakterien einen zyklischen Fluss besitzen kann, um die Zellrespiration bei Abwesenheit von Licht zu gewährleisten (siehe Kapitel 2.2.9). Innerhalb der Rekonstruktion des metabolischen Netzwerkes von *Synechocystis* (siehe Kapitel 2) ist der zyklische Fluss durch den TCA-Zyklus durch zwei metabolische Nebenwege (Bypass) realisiert, die den fehlenden OGDH-Komplex ersetzen.

Tabelle 3.4.: Eine Zusammenfassung der Enzyme, die im zentralen Kohlenstoffstoffwechsel auftreten: Glykolyse und Calvin-Benson-Zyklus. Punkte repräsentieren die Anzahl der Gene, die in den CLOGs mit der jeweiligen metabolischen Funktion vorkommen. Hierbei wird nicht für einzelne CLOGs mit derselben enzymatischen Funktion unterschieden, sondern diese zusammen als ein CLOG aufgeführt (Tabelle aus Beck et al. (2012) [130]).

	Glykolyse											Calvin-Benson-Zyklus					
	GPI	FBP	SBP	FBPA	TPI	PFK	GAPDH	PGM	ENO	PYK	PGK	RPI	TKT	TALDO	PRK	RPE	RuBisCO
Aca11017	••	••	••	••	•	•	••	•••	•	•••	•	••	•	••	•	•	••
Cyn51142	•	•	•	•	•	••	••	•••••	•	•••	•	•	•	••	•	•	••
Cyn8801	•	•	•	•	•	••	••	•••	•	••	•	•	•	••	•	•	••
Glo7421	•	•	••	•	•	-	••	•••••	•	•••	•	••	•	••	•••	•	••
Mic843	•	•	•	•	•	••	••	•••	•	•••	•	•	•••	•	•	•	••
Nos7120	•	•	•	••	•	•	•••	•••	•	••	•	•	••	••	••	•	••
ProMed4	•	-	•	•	•	-	••	••	•	•	•	•	•••	•	•	•	••
Pro9211	•	-	•	•	•	-	••	••	•	•	•	•	•	•	•	•	••
Pro9215	•	-	•	•	•	-	•	••	•	•	•	•	•	•	•	•	••
SycJA23	•	•	•	•	•	•	•••	•••	•	••	•	•	•	•	••	•	••
Syc7002	•	•	•	••	•	•	••	••	•	•	•	•	•	••	••	•	••
Syc7803	•	-	•	••	•	-	•••	••	•	•	•	•	•	•	•	•	••
Syc7942	•	•	•	•	•	•	•••	••••	•	•	•	•	•	•	•	•	••
Syn6803	•	•	•	••	•	••	••	••••	•	••	•	••	•	•	•	•	••
ThermoBP1	•	•	•	•	•	•	••	••	•	••	•	•	•	•	•	•	••
Trich101	•	•	•	••	•	-	•••	••	•	•	•	•	•	••	••	•	••

Eine Möglichkeit besteht darin, dass der Fluss über die drei Reaktionsschritte der Glutamat-Decarboxylase (EC 4.1.1.15), γ -Aminobutyrat-Transaminase (EC 2.6.1.19) und der Succinat-Semialdehyd-Dehydrogenase (EC 1.2.1.16) (SSADH) fließen kann.

Die zweite Möglichkeit beruht darauf, dass seit kurzem auch ein alternativer Nebenweg von Zhang und Bryant (2011) [142] identifiziert wurde. Hierbei handelt es sich um eine neuartige 2-Oxoglutarat-Decarboxylase (EC 4.1.1.71) (2OGDC), die zusammen mit der letztgenannten SSADH, die Konvertierung von 2-Oxoglutarat zu Succinat ermöglicht und damit den TCA-Zyklus schließt [142] (siehe Kapitel 2.2.9). Die zugehörigen metabolischen CLOGs für den zweiten Nebenweg werden von allen Stämmen geteilt außer den drei *Prochlorococcus*-Stämmen sowie *Synechococcus* und *Syc7803*. Die Existenz eines solchen Nebenweges über Succinat-Semialdehyd erklärt auch, dass die Succinat-Thiokinase (EC 6.2.1.5) (STK), welche sonst ein essentielles Enzym innerhalb des TCA-Zyklus darstellt, nur für eine kleine Anzahl von Stämmen annotiert ist. Es wird angenommen, dass in diesen wenigen Fällen die STK dazu dient, vor allem Succinyl-CoA für die Nutzung als Cofaktor im weiteren Stoffwechsel bereitzustellen.

Interessanterweise fehlt in verschiedenen Cyanobakterienstämmen die Succinat-Dehydrogenase (EC 1.3.99.1) (SDH), was die Fähigkeit und Kapazität der zellulären Respiration über den TCA-Zyklus in diesen Stämmen signifikant beeinträchtigt. Hingegen sind CLOGs, welche mit der Fumarat-Hydratase (EC 4.2.1.2) (FH) assoziiert sind, in allen 16 Stämmen vorhanden. Jedoch sind die Gene für *Syc7002* und *Glo7421* nicht als FH annotiert und besitzen keine so hohe Ähnlichkeit mit den Genen wie andere Stämme untereinander, was darauf schließen lässt, dass es sich vielleicht um bifunktionale Enzyme handelt oder die Annotation fehlerhaft ist. An den Ergebnissen aus den FBA (siehe Kapitel 2) ist ersichtlich, dass die FH essentiell für das Recycling

Tabelle 3.5.: Eine Zusammenfassung der Enzyme, die im zentralen Kohlenstoffstoffwechsel beteiligt sind: Der Pentose-Phosphat-Weg (PPP), der Pyruvat-Stoffwechsel und der TCA-Zyklus. Punkte repräsentieren die Anzahl der Gene, die in den CLOGs mit der jeweiligen metabolischen Funktion vorkommen. Hierbei wird nicht für einzelne CLOGs mit derselben enzymatischen Funktion unterschieden, sondern diese zusammen als ein CLOG aufgeführt. Punkte in Klammern entsprechen Gene, die eine andere Annotation besitzen als der zugehörige CLOG, in dem sie aufgeführt sind (Tabelle aus Beck et al. (2012) [130]).

	PPP			Pyruvat-Stoffwechsel					TCA-Zyklus							
	GPD	6PGD	6PGL	PEPC	ME	PPS	PEPK	PDH	CS	ACO	ICD	STK	SDH	FH	MDH	MQO
Aca11017	••	••	•	•	•	•••	-	••	••	•	••	-	••(•)	•	•	-
Cyn51142	•	•	•	•	•	••	•	••	•	•	•	••	••(•)	•	•	-
Cyn8801	•	•	•	•	•	••	•	••	•	•	•	••	••(•)	•	•	-
Glo7421	••	•	•	•	•	••	-	••••	•	•	••	-	••(•)	(•)	•	-
Mic843	•	•	•	•	•	••	•	••	•	•	•	••	••(•)	•	•	-
Nos7120	•	•	•	•	•	••••	-	••	•	•	•	••	••(•)	•	•	-
ProMed4	•	•	•	•	-	-	-	••	•	•	•	-	-	•	-	•
Pro9211	•	•	•	•	-	-	-	••	•	•	•	-	-	•	-	•
Pro9215	•	•	•	•	-	-	-	••	•	•	•	-	-	•	-	•
SycJA23	•	•	•	•	•	•	-	••	•	•	•	-	••	•	•	-
Syc7002	•	•	•	•	•	•	-	••	•	•	•	••	••(•)	(•)	•	-
Syc7803	•	•	•	•	-	-	-	••	•	•	•	-	••	•	-	•
Syc7942	•	•	•	•	•	•	-	••	•	•	•	-	••	•	-	-
Syn6803	•	•	•	•	•	•	-	••	•	•	•	••	••(•)	•	•	-
ThermoBP1	•	•	•	•	•	••	-	••	•	•	•	-	••	•	-	-
Trich101	•	•	•	•	•	••	-	••	•	•	•	••	••	•	-	-

von Fumarat ist, welches als ein obligates Nebenprodukt während des Wachstums entsteht.

Darüber hinaus sind noch weitere Enzyme nur für eine Untergruppe der 16 Stämme annotiert. Insbesondere fehlen das Malatenzym (EC 1.1.1.38) (ME) und die Malat-Dehydrogenase (EC 1.1.1.37) (MDH) in allen α -Cyanobakterien. Bei diesen Stämmen wird die Funktion des Letzteren durch eine Malat:Chinon-Oxidoreduktase (EC 1.1.5.4) (MQO) kompensiert, welche aber nur die irreversible Oxidation von Malat zu Oxalacetat katalysieren kann.

3.3. Schlussfolgerungen

Die rasch wachsende Zahl von vollständig sequenzierten mikrobiellen Genomen bietet neue Möglichkeiten, die mikrobielle Vielfalt in komplexen Umgebungen zu verstehen. In diesem Teil der Arbeit wurde ein besonderer Fokus auf mehrere phototrophe Cyanobakterien gelegt, mit dem Ziel, Erkenntnisse über die Vielfalt des cyanobakteriellen Stoffwechsels zu gewinnen. Cyanobakterien besitzen eine enorme metabolische Diversität und treten in fast allen Umgebungen, wo Licht verfügbar ist, auf. Sie sind daher besonders für eine vergleichende Analyse der genetischen Diversität geeignet.

Die Basis für die vergleichende Analyse war die Definition von Clustern von vermutlich orthologen Genen (*Cluster of Likely Ortholog Genes*) und wie diese in den 16 analysierten Cyanobakterienstämmen verteilt sind. Insgesamt wurden 21.238 unterschiedliche CLOGs in den Cyanobakterienstämmen identifiziert, von denen die Mehrheit (ungefähr 65 %) aus einzelnen Genen besteht und keine vermutlichen Orthologen in einem der anderen analysierten Stämme besitzt. Etwa 3 % der CLOGs sind allen Stämmen zugeordnet und machen somit das Kern-Genom aus, welches von allen 16 Stämmen geteilt wird. Zu beachten ist jedoch, dass dieses Kern-Genom nicht das minimale Set an Genen repräsentiert, welches die Zelle autark lebensfähig macht. Die übrigen CLOGs sind mehreren unterschiedlichen Stämmen zugeordnet, aber nicht in allen 16 Stämmen zu finden. Die Ergebnisse aus dieser Vergleichsanalyse lassen mehrere Schlussfolgerungen zu:

Es gibt keine Hinweise darauf, dass das Pan-Genom der Cyanobakterien beschränkt ist. Vielmehr steigt die Anzahl der gesamten CLOGs mit der Einbeziehung weiterer Cyanobakterienstämmen und zeigt, dass es nicht ausreicht, nur wenige Stämme zu sequenzieren, um die ganze Vielfalt des cyanobakteriellen Genoms zu ergründen und zu erfassen. Während Extrapolationen mit Vorsicht zu betrachten sind [196], geben jedoch die Ergebnisse einen starken Anreiz für neue Sequenzierungsprojekte von weiteren Cyanobakterienstämmen.

Weiterhin unterscheidet sich die Menge der Kern- und einzigartigen CLOGs in Bezug auf die Anreicherung von unterschiedlichen Genannotationen. Kern-CLOGs zeigen eine deutliche Anreicherung von Genen, die gemeinhin mit einer Housekeeping-Funktion assoziiert sind. Dazu gehören unter anderem die Bereiche „Translation“, „DNA-Reparatur“, „Genexpression“, „RNA-Prozessierung und Modifikation“, diverse Transportprozesse sowie zahlreiche Stoffwechsel- und biosynthetische Prozesse. Hingegen umfasst die Annotation von Genen aus den einzigartigen CLOGs eher die Begriffe von näher bestimmbar Funktionen, wie beispielsweise verschiedene regulatorische Prozesse. Bei einer genaueren Betrachtung der Diversität des cyanobakteriellen Stoffwechsels auf Genomebene ist zu erkennen, dass CLOGs, denen eine Stoffwechselfunktion zugewiesen werden kann, innerhalb der Menge der Kern-CLOGs stark überrepräsentiert sind.

Bei einer genaueren Untersuchung der Verteilung dieser metabolischen CLOGs zeigt die Analyse ein sehr vielfältiges Bild in Bezug auf das Vorhandensein von Kern-

Stoffwechselwegen. Mehrere wichtige Stoffwechselwege des zentralen Stoffwechsels sind hoch konserviert, wie der Pentosephosphatweg und der Calvin-Benson-Zyklus. Jedoch sind andere Teile des metabolischen Netzwerkes, insbesondere der Pyruvatstoffwechsel und der TCA-Zyklus, stark fragmentiert.

Im Gegensatz zur herkömmlichen funktionellen Annotation bietet die Annotation von enzymatischen Funktionen den Vorteil, dass der funktionale Kontext des jeweiligen CLOGs, im Hinblick auf den Stoffwechselweg und die benachbarten Reaktionen, mit berücksichtigt werden kann. Deshalb ermöglicht es diese Analyse, die Anwesenheit oder Abwesenheit von bestimmten Enzymen in Bezug auf die metabolische Funktion zu bewerten und liefert dadurch stärkere Kriterien, um fehlerhafte Annotationen oder die Verlässlichkeit der CLOG-Assoziation zu beurteilen. Zum Beispiel schränkt der unvollständig annotierte TCA-Zyklus von Cyanobakterien ihre Fähigkeit, Vorstufenmetaboliten für die Zellatmung zu liefern, stark ein. Dieses Defizit hat Konsequenzen für die funktionale Rolle der anderen Reaktionen, wie unter anderem die der Succinat-Thiokinase (EC 6.2.1.5). Es ist deswegen unwahrscheinlich, dass diese ihre übliche Rolle im TCA-Zyklus übernimmt und sie ist dementsprechend nur für eine geringe Anzahl von Stämmen annotiert.

In diesem Zusammenhang kann auch die Vergleichsanalyse als ersten Schritt der automatisierten Netzwerkrekonstruktion gesehen werden. Genomskalige Modelle des Zellstoffwechsels werden immer wichtiger für eine Vielzahl von biotechnologischen Anwendungen, sind aber derzeit oft auf eine kleine Anzahl von Modellstämmen beschränkt [112]. Die Rekonstruktionen von multiplen Stämmen können erheblich von einer gründlichen Analyse der metabolischen Diversität profitieren, die unter der Menge von bereits sequenzierten Cyanobakterien existiert. Ein integrativer Prozess von der Genomanalyse bis zur Modellrekonstruktion wird schließlich zu einem besseren Verständnis der metabolischen und ökologischen Funktionen von Bakterienspezies führen.

**Das metabolische Netzwerk von
Synechococcus elongatus PCC 7942**

Vorwort

Mit den gewonnenen Erkenntnissen über die Diversität des cyanobakteriellen Stoffwechsels aus Kapitel 3, ergibt sich die Frage, wie das komplette metabolische Netzwerk eines zweiten Cyanobakteriums aufgebaut wäre und welche Resultate eine Modellierung im Vergleich zum Netzwerk von *Synechocystis* sp. PCC 6803 liefern würde.

Mit dieser Motivation wird im nachfolgenden Kapitel die Rekonstruktion und Modellierung eines genomskaligen metabolischen Netzwerkes von *Synechococcus elongatus* PCC 7942 beschrieben, eines weiteren Modellorganismus unter den Cyanobakterien. Das vorgestellte Netzwerk besitzt im Vergleich zum Netzwerk von *Synechocystis* sp. PCC 6803 einen etwas geringeren Umfang und damit einhergehend eine geringere metabolische Kapazität und Variabilität. Insbesondere werden die Unterschiede im Zentralstoffwechsel hervorgehoben, erläutert und deren Auswirkungen auf die maximale Wachstumsrate und die optimale Flussverteilung im Detail untersucht.

Am Ende des Kapitels wird abschließend eine Verifikation des Netzwerkes von *Synechococcus elongatus* PCC 7942 mithilfe von Proteomikdaten vorgestellt, die von der Arbeitsgruppe von Ilka Maria Axmann zur Verfügung gestellt wurden. Demnach konnten die Translationsprodukte von ungefähr 89 % der im metabolischen Netzwerk assoziierten Gen-Protein-Verknüpfungen detektiert werden.

4.1. Material und Methoden

4.1.1. Rekonstruktion des metabolischen Netzwerkes von *Synechococcus elongatus* PCC 7942

Für die Rekonstruktion des metabolischen Netzwerkes von *Synechococcus elongatus* PCC 7942 (*Synechococcus*) wurden zunächst alle Gene, die im vorher rekonstruierten Netzwerk von *Synechocystis* sp. PCC 6803 (*Synechocystis*) vorhanden sind, in eine Liste übertragen. Anschließend wurde für jedes Gen in dieser Liste geprüft, ob ein korrespondierendes Gen im Genom von *Synechococcus* existiert. Zu diesem Zweck wurde untersucht, ob in dem jeweiligen CLOG (siehe Kapitel 3), in dem das *Synechocystis*-Gen vorkommt, ein Gen für *Synechococcus* vorhanden ist. War dies der Fall, wurde das entsprechende *Synechococcus*-Gen mit den zugehörigen Reaktionen aus dem *Synechocystis*-Netzwerk in das *Synechococcus*-Netzwerk übertragen. Für alle übrigen *Synechococcus*-Gene, die in CLOGs vorhanden waren, denen eine enzymatische Funktion zugewiesen wurde, wurde geprüft, ob deren entsprechenden metabolischen Reaktionen mit schon vorhanden Metaboliten aus dem *Synechococcus*-Netzwerk interagierten und dann anschließend in das Netzwerk übernommen. Sofern dies nicht der Fall war, wurde untersucht, ob durch mehrere annotierte Enzyme beziehungsweise deren Reaktionsschritte, die spezifische Lücke im Netzwerk geschlossen werden konnte. Falls dies möglich war, wurden alle entsprechenden Gene und deren assoziierten Reaktionen ins Netzwerk integriert.

Schlussendlich wurden noch alle nicht durch Enzyme katalysierten Reaktionen aus dem *Synechocystis*-Netzwerk in das *Synechococcus*-Netzwerk transferiert. Ferner wurden die Kompartimentierungen und die Zuweisung der Reaktionen zu diesen Kompartimenten für die Rekonstruktion des *Synechococcus*-Netzwerkes übernommen. Als finalen Schritt wurde die Biomassezielfunktion aus dem *Synechocystis*-Netzwerk an die Gegebenheiten des Stoffwechsels von *Synechococcus* angepasst: Da *Synechococcus* nicht in der Lage ist, irgendeine Variante von Tocopherol zu bilden, wurde dieses vollständig aus der Biomasseformel entfernt und deren Anteil zu den anderen Bestandteilen der Pigmentkomponente addiert. Desgleichen sind noch weitere Pigmente in *Synechococcus* nicht abundant [219]. So wurde auch der Anteil von Echinenon zum Anteil von Zeaxanthin addiert. Der Anteil von Succinyl-CoA in der Biomasseformel wurde durch Acetyl-CoA ersetzt. Aufgrund der fehlenden Δ^6 - und Δ^{15} -Fettsäuredesaturasen im Genom von *Synechococcus* [212] wurden die mehrfach ungesättigten Membranlipide mit mehr als zwei Doppelbindungen nicht berücksichtigt und deren prozentualer Faktor zu dem von Linolsäure addiert.

4.1.2. Flussbilanzanalyse des metabolischen Netzwerkes von *Synechococcus elongatus* PCC 7942

Für die Simulation des phototrophen Wachstums und des Wachstums bei Abwesenheit von Licht von *Synechococcus* wurden die gleichen Bedingungen und Parameter wie beim Netzwerk von *Synechocystis* sp. PCC 6803 festgelegt. Alle Eingangsparameter der FBA des *Synechococcus*-Modells für das phototrophe Wachstum sowie das Wachstum bei Abwesenheit von Licht, sind in der Tabelle 4.1 aufgeführt.

Tabelle 4.1.: Eingangsparameter für die FBA des metabolischen Netzwerkes von *Synechococcus elongatus* PCC 7942. Aufgeführt sind die festgelegten Flussraten [$\text{mmol gDW}^{-1} \text{h}^{-1}$] für die entsprechenden Reaktionen im Netzwerk für das phototrophe Wachstum (Licht) und bei Abwesenheit von Licht (Dunkel).

Reaktion	Reaktionsnummer	Licht	Dunkel
Lichteingang	PR0001	18,7	0
Diverse Kohlenstoff-transporter	TR0006	0	0
	TR0008	0	0
	TR0009	0	0
	TR0011	0	0
	TR0014	0	0
	TR0014_2	0	0
	TR0034	0	0
	TR0035	0	0
basaler ATP-Verbrauch	GE0001	0,13	0,07
Cytochrom-c-Oxidase	PR0010	0,2263	0
Mehler-ähnliche Reaktion	PR0033	0,2263	0
Sauerstoffradikale am PS II	PR0034	0,0477	0
Mehler-Reaktion	PR0032	0,0473	0
Glykogenmobilisierung	GLYC	0	0,01
Biomassezielfunktion	BM0009	max.	max.

4.2. Ergebnisse und Diskussion

4.2.1. Das metabolische Netzwerk von *Synechococcus elongatus* PCC 7942

Neben *Synechocystis* sp. PCC 6803 zählt auch *Synechococcus elongatus* PCC 7942 zu den Modellorganismen unter den Cyanobakterien. Insbesondere für den Bereich der Erforschung der circadianen Uhr in Prokaryoten inklusive deren Mechanismus existieren viele Studien und Analysen [29, 220–222]. Einen sehr guten Ausgangspunkt für die Rekonstruktion eines metabolischen Netzwerkes von *Synechococcus* liefert die Vergleichsanalyse aus Kapitel 3. Unter Miteinbeziehung des metabolischen Netzwerkes von *Synechocystis*, welches als Basis diente, wurde eine genomskalige Netzwerkrekonstruktion des Stoffwechsels für *Synechococcus* erstellt. Fördernd für diesen Prozess der Rekonstruktion wirkten hierbei die Erfahrungen und Gegebenheiten, die aus der schon weit fortgeschrittenen Rekonstruktion des *Synechocystis*-Netzwerkes übertragen werden konnten.

Für die Rekonstruktion von *Synechococcus* wurden die gleichen Stoffwechselwege entsprechend dem *Synechocystis*-Netzwerk betrachtet und in ein Kernnetzwerk integriert. Entsprechend dem *Synechocystis*-Netzwerk wurde für diesen Prozess beachtet, dass eine Synthese aller Biomassekomponenten, Cofaktoren und Metaboliten im Netzwerk ermöglicht wurde. Die aktuelle Netzwerkrekonstruktion von *Synechococcus* umfasst zurzeit 581 Gene, die für 468 Enzyme oder Enzymkomplexe kodieren. Diese resultieren in 666 zusammenhängenden Stoffwechselreaktionen, welche 574 Stoffwechselmetaboliten umsetzen. Somit sind es 94 Reaktionen weniger als im *Synechocystis*-Netzwerk. Die Reaktionen sind sieben unterschiedlichen Kompartimenten zugeteilt, welche analog zum *Synechocystis*-Netzwerk spezifiziert wurden. Da *Synechococcus* phototroph wächst, beschränkten sich die Simulationen ausschließlich auf das phototrophe Wachstum und die Nutzung von Glykogen als Nachtspeicher bei Abwesenheit von Licht. Entsprechend dem vorher rekonstruierten und analysierten Netzwerk von *Synechocystis* wurde die Aufnahme und die Verwertung von zusätzlichen komplexeren organischen Kohlenstoffquellen nicht in Betracht gezogen. Die Rekonstruktion des *Synechococcus*-Netzwerkes ist im elektronischen Anhang dieser Arbeit als SBML- und Excel-Datei verfügbar.

4.2.2. Der Zentralstoffwechsel von *Synechococcus elongatus* PCC 7942 im Vergleich zu *Synechocystis* sp. PCC 6803

Synechococcus elongatus PCC 7942 und *Synechocystis* sp. PCC 6803 gehören beide zu den β -Cyanobakterien und wurden zuerst aus Süßwasserhabitaten isoliert. Beide Stämme leben als Einzelzellen und sind nicht in der Lage, elementaren Stickstoff zu fixieren (siehe Tabelle 3.1). Betrachtet man die Genomgröße, ausgenommen der Plasmide, ist hingegen auszumachen, dass das Genom von *Synechocystis* um mehr als 25 % größer ist im Vergleich zu *Synechococcus* (3,57 Mb zu 2,8 Mb, siehe Tabelle 3.1).

Dieser Sachverhalt spiegelt sich auch auf der Ebene der metabolischen Kapazität wieder. Während *Synechocystis* sehr variabel ist und auf mindestens drei verschiede-

ne Speicherverbindungen zurückgreifen kann, dazu gehören unter anderem Glykogen, Poly- β -Hydroxybutyrat (PHB) und Cyanophycin, sind im Genom von *Synechococcus* nur Gene für die Glykogen-Synthese und –Mobilisierung zu finden. Gene für die Synthese von PHB oder Cyanophycin sind nicht vorhanden (siehe Kapitel 3.2.6). Außerdem ist *Synechococcus* nicht in der Lage, die verschiedenen Tocopheroltypen zu bilden, welche von *Synechocystis* zum Schutz gegen oxidativen Stress synthetisiert werden können [223]. Zudem kann das Carotinoid Echinenon nicht von *Synechococcus* gebildet werden [219].

Bei genauerer Betrachtung des Fettstoffwechsels fällt auf, dass im Genom von *Synechococcus* keine Δ^6 - und Δ^{15} -Fettsäuredesaturasen annotiert sind [212] und mehrfach ungesättigte Membranlipide mit mehr als zwei Doppelbindungen im Stoffwechselnetzwerk nicht synthetisiert werden können. Darunter fallen vor allem γ -Linoleoyl-ACP, α -Linoleoyl-ACP und Stearidonoyl-ACP. Ergänzend fehlt in *Synechococcus* der gesamte Harnstoffmetabolismus, was zur Folge hat, dass jegliches Enzym, welches in irgendeiner Weise Harnstoff bildet oder umsetzt, nicht in *Synechococcus* vorkommt. Dies betrifft unter anderem die Synthese von Putrescin und folglich auch die von Spermidin. Putrescin wird in *Synechocystis* über eine Arginin-Decarboxylase (EC 4.1.1.19) (slr0662 und slr1312) und eine Agmatinase (EC 3.5.3.11) (sll0228 oder sll1077) synthetisiert. Die Arginin-Decarboxylase decarboxyliert Arginin zu Agmatin, welches wiederum durch die Agmatinase zu Putrescin und zu Harnstoff hydrolysiert wird. Da bei *Synechococcus* keine Umsetzung von Harnstoff im Stoffwechsel möglich ist, wird der letzte Schritt der Putrescin-Synthese, also die Agmatinase, durch Kombination einer Agmatin-Deiminase (EC 3.5.3.12) (Synpcc7942_2402 oder Synpcc7942_2461, durch eine BLAST-Suche identifiziert), welche Agmatine zu N-Carbamoylputrescin und Ammonium hydrolysiert, mit einer N-Carbamoylputrescin-Amidohydrolase (EC 3.5.1.53) (Synpcc7942_2145), die anschließend aus dem N-Carbamoylputrescin Putrescin bildet, kompensiert.

Bei der Synthese der verzweigt-kettigen Aminosäuren ist auszumachen, dass *Synechococcus* im Gegensatz zu *Synechocystis* nur die Möglichkeit aufweist, 2-Oxobutyrat über dessen alternativen Syntheseweg herzustellen, der von Wu et al. (2010) [224] beschrieben wurde. *Synechocystis* besitzt hingegen zusätzlich noch den weitverbreiteten Syntheseweg über eine Threonin-Deaminase (EC 4.3.1.19) (slr2072), die in *Synechococcus* fehlt. Des Weiteren ist es in *Synechococcus* nicht möglich, Glycerat, welches unter anderem als Produkt der Photorespiration entstehen kann [119], über Glycerol zu Glycerol-1-Phosphat abzubauen, um es von dort weiter im Stoffwechsel wiederzuverwerten. Bei *Synechococcus* existiert nur der Stoffwechselweg über die direkte Konversion von Glycerat zu Glycerat-3-Phosphat.

Ein gravierender Unterschied im Stoffwechsel von *Synechococcus* zu *Synechocystis* wird bei einer näheren Betrachtung des TCA-Zyklus inklusive der angrenzenden Reaktionen erkennbar (siehe Abbildung 4.1). Von den konventionellen Enzymen des TCA-Zyklus sind im Vergleich zu *Synechocystis* (Abbildung 4.1 A) keine Gene, die für eine Succinat-Thiokinase (EC 6.2.1.5) (STK) oder Malat-Dehydrogenase (EC 1.1.1.37) (MDH) kodieren, zu finden. Insbesondere letztere, welche bei den α -Cyanobakterien

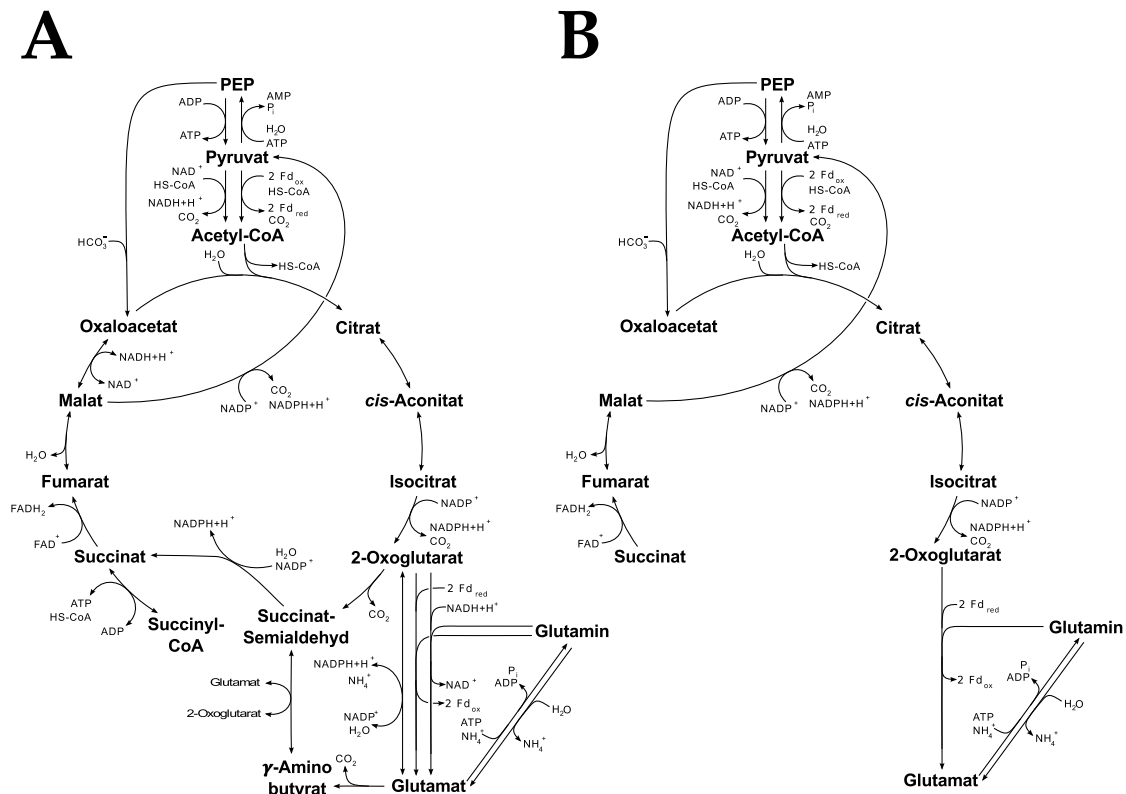


Abbildung 4.1.: Unterschiede auf Ebene des TCA-Zyklus zwischen A) *Synechocystis* sp. PCC 6803 und B) *Synechococcus elongatus* PCC 7942. Während es bei A) *Synechocystis* möglich ist, auf zwei Weisen einen zyklischen Fluss zu gewährleisten, ist laut Genannotation bei B) *Synechococcus* ein zyklischer Fluss durch den TCA-Zyklus nicht möglich. Zudem fehlt eine Malat-Dehydrogenase, wodurch nur ein Fluss über das CO₂-freisetzende Malatenzym möglich ist.

durch eine Malat:Chinon-Oxidoreduktase (EC 1.1.5.4) (MQO) kompensiert wird (siehe Kapitel 3.2.7), wird durch keine entsprechende Reaktion ausgeglichen. Demzufolge ist ein Recycling von Malat beziehungsweise Fumarat, welches während unterschiedlicher Synthesewege als Nebenprodukt entstehen kann, nur über das Kohlenstoffdioxid-freisetzende Malatenzym (EC 1.1.1.38) (ME) möglich. Der resultierende Netto-Energiegewinn mittels eines zyklischen Flusses durch den TCA-Zyklus ist dadurch stark eingeschränkt.

Das Szenario eines zyklischen Flusses durch den TCA-Zyklus ist jedoch laut Genomannotation ohnehin nicht realisierbar, da in *Synechococcus* weder Enzyme für den GABA-Shunt [93] noch für eine 2-Oxoglutarat-Decarboxylase (EC 4.1.1.71) [142] zu finden sind (siehe Abbildung 4.1 B). Beide Stoffwechselwege würden gewährleisten, dass der fehlende 2-Oxoglutarat-Dehydrogenase-Komplex bei Cyanobakterien kompensiert werden könnte. Auch ist kein Gen im Genom von *Synechococcus* für eine Glutamat-Dehydrogenase (EC 1.4.1.4) (GDH) annotiert, welche direkt an den TCA-Zyklus an-

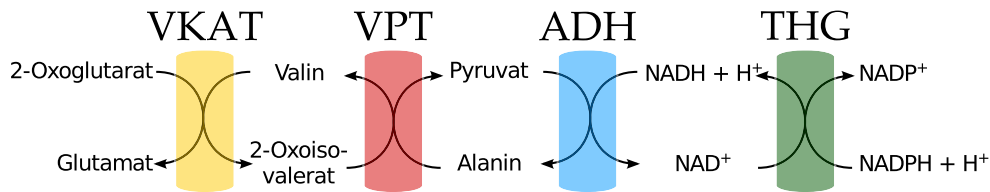


Abbildung 4.2.: Alternativer Stoffwechselweg zur Glutamat-Dehydrogenase in *Synechococcus elongatus* PCC 7942. Involviert sind dabei die VKAT: Verzweigtkettige-Aminosäuren-Transaminase, die VPT: Valin-Pyruvat-Transaminase, die ADH: Alanin-Dehydrogenase und die THG: Transhydrogenase.

schließt. Dieses wichtige Enzym des Glutamat-Stoffwechsels kann jedoch stöchiometrisch äquivalent durch ein Zusammenspiel von drei Reaktionen im Netzwerk ersetzt werden. Hierzu gehören die Alanin-Dehydrogenase (EC 1.4.1.1) (Synpcc7942_1760), die Valin-Pyruvat-Transaminase (EC 2.6.1.66) (Synpcc7942_2258) und die nicht spezifische Verzweigtkettige-Aminosäuren-Transaminase (EC 2.6.1.42) (Synpcc7942_1029) (siehe Abbildung 4.2). Ob ein solcher Zyklus thermodynamisch realisierbar ist und *in vivo* abläuft, müsste noch weiter untersucht werden.

Im Gegensatz zu *Synechocystis* sind die essentiellen Gene, die für die Methioninsynthese benötigt werden, bei *Synechococcus* annotiert. Die Synthese kann über die Reaktionsschritte der Homoserin-O-Acetyltransferase (EC 2.3.1.31) (Synpcc7942_1714) und die O-Acetylhomoserinlyase (EC 2.5.1.49) (Synpcc7942_0370) verlaufen.

4.2.3. Flussbilanzanalyse des metabolischen Netzwerkes von *Synechococcus elongatus* PCC 7942

Bei der Analyse des Netzwerkes von *Synechocystis* sp. PCC 6803 hatte sich die Flussbilanzanalyse (*Flux Balance Analysis*) (FBA) als probates Hilfsmittel bewährt, um zuverlässige Aussagen über die optimale Flussverteilung und Wachstumsraten zu treffen [93, 94]. Für die Simulation des phototrophen Wachstums bei *Synechococcus elongatus* PCC 7942 wurden die gleichen Bedingungen und Einschränkungen wie beim *Synechocystis*-Netzwerk festgelegt.

Ähnlich den Simulationen für *Synechocystis* ist die resultierende Lösung der optimalen Flussverteilung für das phototrophe Wachstum nicht eindeutig und es herrscht erneut Variabilität, vor allem bei den Reaktionen des Calvin-Benson-Zyklus (siehe Abbildung 4.3). Der größte Unterschied besteht jedoch im Bereich des TCA-Zyklus. Auf Grund der fehlenden Malat-Dehydrogenase (EC 1.1.1.37) trägt die CO₂-freisetzende Reaktion des Malatenzyms (EC 1.1.1.38), im Gegensatz zu der optimalen Flussverteilung von *Synechocystis*, per se einen Fluss. Eine der höchsten Flussraten außerhalb des Calvin-Benson-Zyklus besitzt die Reaktion der Alanin-Dehydrogenase. Wie im vorangegangenen Kapitel beschrieben, fehlen in *Synechococcus* die Gene für eine Glutamat-Dehydrogenase, die Glutamat für eine große Anzahl an Stoffwechselreaktionen bereitstellt. Dies hat zur Folge, dass deren Funktion durch die Alanin-Dehydrogenase zusammen mit drei weiteren Reaktionsschritten bei den Simulationen

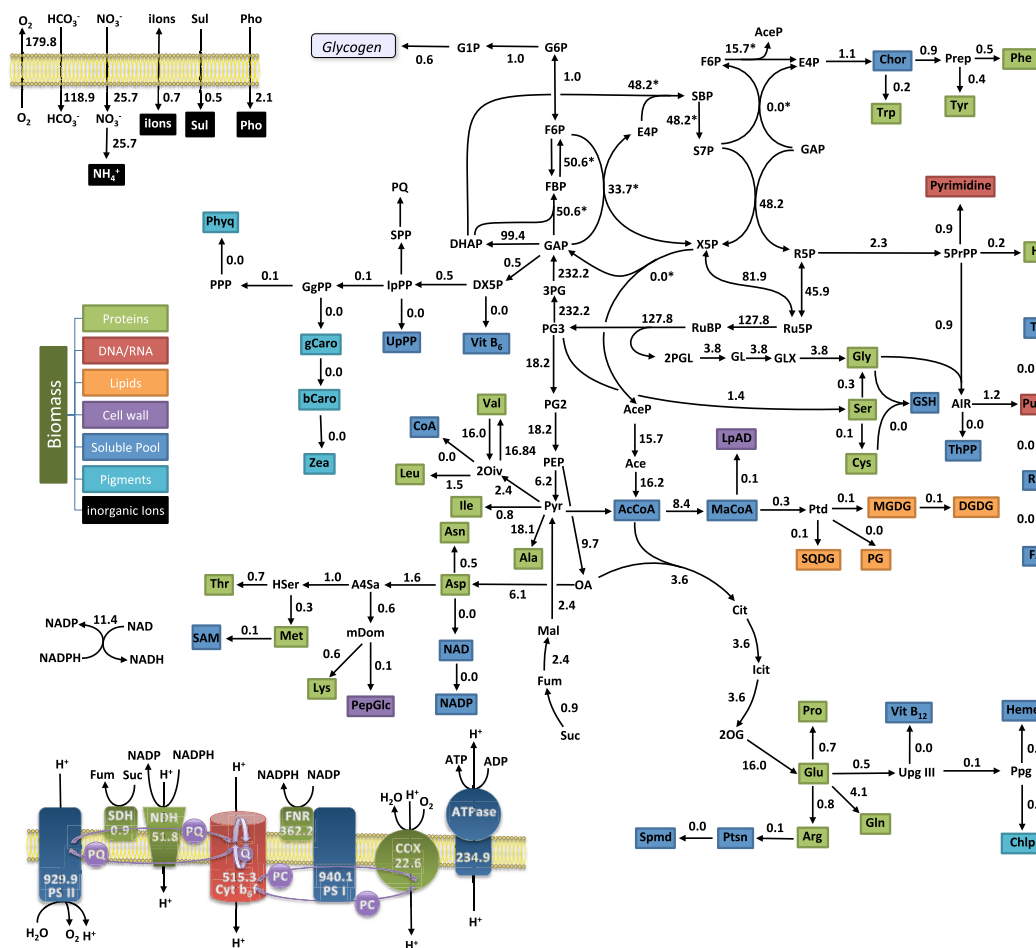


Abbildung 4.3.: Optimale Flussverteilung für *Synechococcus elongatus* PCC 7942 unter phototrophen Bedingungen. Für die Flussverteilung wurde auf eine maximale Biomasseproduktion optimiert. Die Flüsse sind in $10^{-2} \text{ mmol gDW}^{-1} \text{ h}^{-1}$ gegeben. Variable Flüsse sind durch * markiert. Alle Reaktionen mit einer Flussrate > 0 aber $< 0,05 \cdot 10^{-2} \text{ mmol gDW}^{-1} \text{ h}^{-1}$ sind durch einen Fluss von 0,0 gekennzeichnet. Die Erklärung der Abkürzungen für die verwendeten Metaboliten sind im Anhang in der Tabelle A.3 aufgeführt.

kompensiert wird (siehe Abbildung 4.2), wodurch sich der hohe Flussdurchsatz erklären lässt.

Zusammengefasst lässt sich für das Netzwerk von *Synechococcus elongatus* PCC 7942 bei einer Simulation des phototrophen Wachstums die Aussage formulieren, dass das Netzwerk nicht die Kapazität und Effektivität des Synechocystis-Netzwerkes besitzt, um die einzelnen Komponenten der Biomasse zu synthetisieren. Denn obwohl die Biomasseformel für das *Synechococcus*-Netzwerk auf Grund der unter anderem fehlenden Desaturasen und Tocopherole (siehe Kapitel 4.2.2) vereinfacht wurde, erhält man unter den gleichen Simulationsbedingungen mit denselben Eingangsparametern (Lichteinfall und Transportvorgänge), eine geringere Wachstumsrate im Vergleich zum Synechocystis-Netzwerk. Dieser Unterschied ist allerdings nur minimal

und beträgt ($0,0287 \text{ h}^{-1}$ für *Synechococcus* zu $0,0288 \text{ h}^{-1}$ für *Synechocystis*). Die optimale Flussverteilung für die Simulation des phototrophen Wachstums von *Synechococcus* ist in Abbildung 4.3 gezeigt.

Bei der Simulation des Wachstums bei Nacht beziehungsweise Abwesenheit von Licht gibt es eine wesentlichere Umverteilung der optimalen Flusslösung im Gegensatz zu *Synechocystis*. Bei Abwesenheit von Licht existiert kein zyklischer Fluss durch den TCA-Zyklus, da wie schon im vorangegangenen Kapitel erwähnt, dieser nicht durch einen der bekannten Wege (Bypass) geschlossen werden kann.

Die Reduktionsäquivalente, welche hauptsächlich in Form von NADPH gebraucht werden, werden über den oxidativen Pentosephosphatweg (OPP) bereitgestellt. Diese Lösung ist nicht so effektiv wie der zyklische Fluss durch den TCA-Zyklus und resultiert dementsprechend in einer geringeren Kapazität beziehungsweise Biomasseproduktion im Vergleich zum Netzwerk von *Synechocystis* (siehe Tabelle 4.2). Auch die Fähigkeit Energie in Form von ATP, welches für die Erhaltung der Zellfunktionen benötigt wird, kann prinzipiell nicht in dem Maße und so effektiv synthetisiert werden, wie es im *Synechocystis*-Netzwerk der Fall ist. Die optimale Flussverteilung für das Wachstum bei Abwesenheit von Licht ist in Abbildung 4.4 dargestellt.

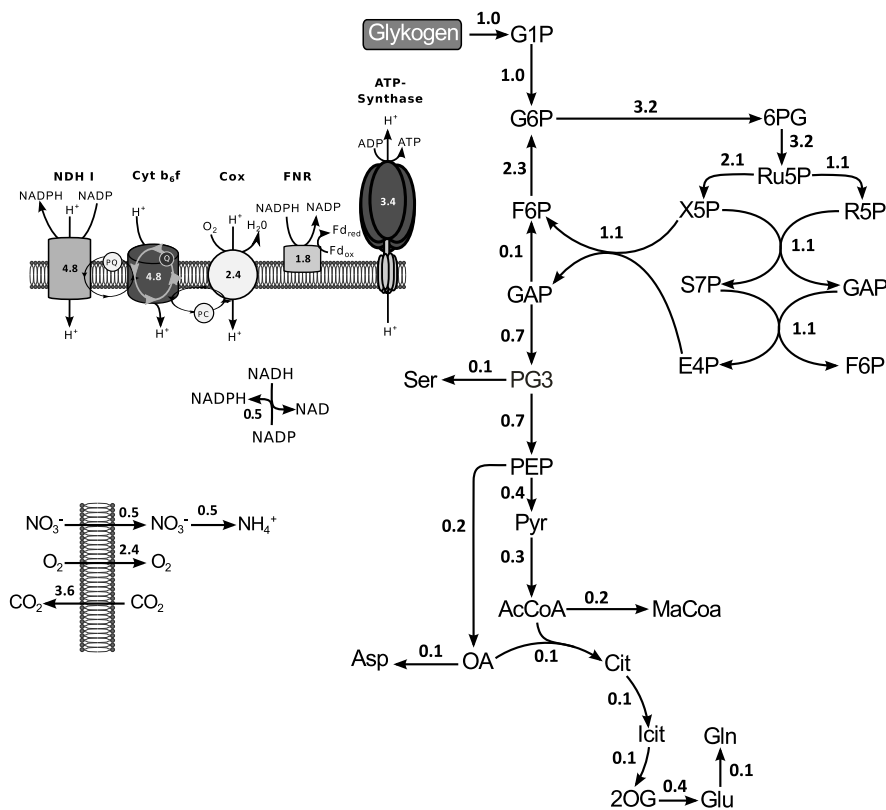


Abbildung 4.4.: Ausschnitt des Zentralstoffwechsels von der optimalen Flussverteilung für *Synechococcus elongatus* PCC 7942 bei Abwesenheit von Licht. Die Flüsse sind in $10^{-2} \text{ mmol gDW}^{-1} \text{ h}^{-1}$ gegeben. Die Erklärung der Abkürzungen für die verwendeten Metaboliten sind im Anhang in der Tabelle A.3 aufgeführt.

Tabelle 4.2.: Kernmerkmale des metabolischen Netzwerkes von *Synechocystis* sp. PCC 6803 im Vergleich zum Netzwerk von *Synechococcus elongatus* PCC 7942.

	<i>Synechocystis</i> sp. PCC 6803	<i>Synechococcus</i> <i>elongatus</i> PCC 7942
Anzahl der Reaktionen	759	666
Anzahl der Metaboliten	601	574
Anzahl der Gene	677	589
max. phototrophe Wachstumsrate [h^{-1}]	0,0288	0,0287
max. Restwachstumsrate bei Abwesenheit von Licht [h^{-1}]	$0,585 \cdot 10^{-3}$	$0,573 \cdot 10^{-3}$
zyklischer Fluss über den TCA- Zyklus bei Abwesenheit von Licht	•	-
Fluss über den OPP bei Abwe- senheit von Licht	-	•

4.2.4. Verifizierung des metabolischen Netzwerkes von *Synechococcus elongatus* PCC 7942

Einer der wichtigsten Schritte, um die Glaubwürdigkeit eines metabolischen Modells zu unterstützen, besteht darin, den Sachverhalt zu überprüfen, ob wirklich alle Reaktionsschritte, die in der jeweiligen Rekonstruktion des Stoffwechsels aufgeführt sind, auch *in vivo* in der Zelle existieren. Ein Augenmerk liegt dabei natürlich auf der Verifizierung jedes einzelnen Enzyms, ob dieses zu einem bestimmten Zeitpunkt in der Zelle exprimiert wird und aktiv ist und somit die gewünschte Reaktion katalysieren kann. Leider sind solche Daten in der Praxis meist nicht verfügbar und wenn, dann nur für einzelne Reaktionsschritte vorhanden. Eine grobe Annäherung hierfür liefert die Tatsache, ob überhaupt zu irgendeinem Zeitpunkt oder unter irgendeiner Bedingung die Translation der entsprechenden Gene zu Proteinen erfolgt. Solche entsprechenden Daten werden durch Proteomik-Analysen unter Verwendung eines Massenspektrometers erzeugt. Im Laufe dieser Arbeit war es möglich, einen solchen Datensatz für das Cyanobakterium *Synechococcus* von der Arbeitsgruppe von Ilka Maria Axmann auszuwerten (unveröffentlichte Daten [225]). Dieser Datensatz beschreibt unterschiedliche Proteinkonzentrationen über den Verlauf mehrerer Tagesperioden. Für den Abgleich mit dem Modell des Stoffwechselnetzwerkes von *Synechococcus* wurden alle Proteinstrukturen und -fragmente, die jemals zu einem der unterschiedlichen Zeitpunkte gemessen werden konnten, ermittelt und überprüft, ob diese mit metabolischen Enzymen aus dem Netzwerk übereinstimmen.

Das metabolische Netzwerk von *Synechococcus* beinhaltet 589 unterschiedliche Gene (siehe Tabelle 4.2). Von diesen 589 metabolischen Genen konnte die Expression von 526 durch die Detektion des jeweiligen Proteinproduktes bestätigt werden. Dies entspricht mehr als 89 % aller Gene im Netzwerk. Demgegenüber stehen 63 verbleibende Gene, die entweder unter den Bedingungen der Probenentnahme für die Proteinmessungen nicht exprimiert und translatiert wurden oder deren Proteinprodukte nicht durch die Analysemethoden detektierbar waren. Von diesen 63 verbleibenden Genen sind neun Gene als Isoenzyme zu identifizieren, das heißt, ihre assoziierte enzymatische Reaktion ist im metabolischen Netzwerk mindestens zweimal vorhanden. Ein Beispiel für diesen Fall sind die beiden *gap1*-Gene (Synpcc7942_0245 und Synpcc7942_1939), von denen aber nur das Translationsprodukt eines Gens (Synpcc7942_0245) in den Proteindaten identifiziert werden konnte. Die übrigen 54 nicht detektierten Proteine umfassen 17 Proteine beziehungsweise Proteinkomponenten, die als Transportproteine fungieren, 17 Proteine oder Untereinheiten sind direkt in der Atmungskette beziehungsweise am Photosyntheseapparat involviert, sechs in den Synthesen der Fettsäuren, vier in den Pigmentsynthesen und zwei bei der Peptidoglycansynthese. Die restlichen acht Proteine sind über unterschiedliche Stoffwechselwege verteilt.

4.3. Schlussfolgerungen

Der Vergleich unterschiedlicher Organismen in Hinsicht auf ihre metabolische Kapazität ist ein wichtiger Aspekt, um profunde Aussagen über die Eignung des Organismus zum Beispiel für die Anwendung in der Biotechnologie, zu treffen. Unter diesem Gesichtspunkt ist auch die Stammauswahl bei den unterschiedlichen Cyanobakterienstämmen, die für eine Produktion verschiedenster Biokraftstoffe in Frage kommen können, zu sehen.

In dieser Hinsicht wurde im Laufe dieser Arbeit ein genomskaliges Netzwerk des Stoffwechsels von *Synechococcus elongatus* PCC 7942 rekonstruiert. Unter Berücksichtigung des schon vorhandenen Netzwerkes des Cyanobakteriums *Synechocystis* sp. PCC 6803 und der vergleichenden Genomanalyse der 16 verschiedenen Cyanobakterienstämmen auf Ebene des Stoffwechsels, konnten diese Informationen mit in die Rekonstruktion einfließen. Das resultierende Netzwerk besitzt, auf die Anzahl der unterschiedlichen Stoffwechselwege bezogen, nahezu den gleichen Umfang wie das zuvor rekonstruierte *Synechocystis*-Netzwerk (siehe Kapitel 2). Jedoch ist die Gesamtanzahl der involvierten Stoffwechselreaktionen und –schritte im Vergleich zu *Synechocystis* geringer, welches im Einklang mit der geringeren Genomgröße von *Synechococcus* (3,57 Mb zu 2,8 Mb) steht.

In der Tat hat das Nichtvorhandensein einiger Stoffwechselreaktionen wesentlichen Einfluss auf die maximale Wachstumsrate beziehungsweise die Kapazität und Effektivität, bestimmte Biomassekomponenten zu synthetisieren. Beim phototrophen Wachstum hat das Fehlen der Threonin-Deaminase (4.3.1.19) und der Fluss über das kohlenstoffdioxidfreisetzende Malatenzym minimale Auswirkungen auf die maximale Wachstumsrate. Zwar hat die Defizienz, das durch die Photorespiration zwangsweise auftretende Glyoxylate nicht über Glycerat und Glycerol zu *sn*-Glycerol-1-Phosphat abbauen zu können sowie das Nichtvorhandensein eines Harnstoffwechsels, keinen direkten Einfluss auf die durch die FBA errechnete optimale Flussverteilung. Jedoch hat dieser Sachverhalt eine nicht so hohe metabolische Variabilität und Anpassungsfähigkeit zur Folge, welche unter anderen oder nicht optimalen Wachstumsbedingungen, signifikanteren Einfluss auf das Wachstum nehmen könnte.

Bei einer genaueren Betrachtung der metabolischen Kapazitäten des zentralen Stoffwechsels von den beiden Cyanobakterien ist zu erkennen, dass *Synechococcus* in einigen Schlüsselreaktionen nicht die metabolischen Charakteristika aufweist, die vor allem für ein optimales Wachstum bei Abwesenheit von Licht beziehungsweise beim heterotrophen Wachstum benötigt werden. Zweifelsohne wird vorwiegend die Stoffwechselkapazität in größerem Maße dadurch beschränkt, dass die Mobilisierung von Glykogen, welches als Kohlenstoffspeicherverbindung für Zeiten bei Abwesenheit von Licht dient, nicht über einen zyklischen Fluss durch den TCA-Zyklus laufen kann, sondern die Gewinnung von Reduktionsäquivalenten ausnahmslos über den Pentosephosphatweg ablaufen muss. Dieser Befund ist zum Teil kontrovers zu Aussagen von Zhang und Bryant (2011) [142], die proklamieren, dass bei allen β -Cyanobakterien der TCA-Zyklus durch deren neu identifizierte 2-Oxoglutarat-Decarboxylase geschlossen

werden kann und somit einen zyklischen Fluss ermöglicht. Da jedoch für *Synechococcus* weder eine 2-Oxoglutarat-Decarboxylase noch eine Succinat-Semialdehyd-Dehydrogenase und eine Malat-Dehydrogenase im Genom identifiziert werden können, bleibt der TCA-Zyklus sehr fragmentär.

Unter diesem Gesichtspunkt kann auch der Sachverhalt betrachtet werden, dass *Synechococcus* nur eine der drei untersuchten Hauptspeicherverbindungen bei Cyanobakterien, nämlich Glykogen, aufweist. Zudem gibt dieser Aspekt Anhaltspunkte darauf, dass *Synechococcus* nicht die große Variabilität und Kompetenz des heterotrophen Wachstums von anderen Cyanobakterien, wie zum Beispiel *Synechocystis*, besitzt. Da jedoch in der Analyse dieser Arbeit andere Speicherstoffverbindungen außer Glykogen beim Wachstum unter Abwesenheit von Licht nicht berücksichtigt wurden, können sich keine konkreten oder direkten Aussagen daraus formulieren lassen.

Die reale Situation in der Zelle, also *in vivo*, überschneidet sich nicht immer im vollen Maße mit den Ergebnissen, die die mathematischen Modellierungen biologischer Probleme liefern. Um die Aussagekraft des mathematischen genomskaligen Modells zu festigen, ist die Verknüpfung einer systematischen Analyse mit Hochdurchsatzdaten essentiell. In Fall des Stoffwechselmodells von *Synechococcus* wurde die resultierende Rekonstruktion des Stoffwechsels mit Proteomdatensätzen von *in vivo* Messungen zu unterschiedlichen Zeitpunkten eines Tagesablaufs abgeglichen [225]. Die Detektion aller Proteine beziehungsweise Proteinfragmente zu jedem beliebigen Zeitpunkt wurde mit den im metabolischen Netzwerk auftretenden Stoffwechselenzymen verglichen. Es konnte gezeigt werden, dass ca. 89 % aller im Netzwerk agierenden Enzyme und Enzymuntereinheiten und somit fast das gesamte Netzwerk, durch die Proteinmessungen abgedeckt werden. Unter den fehlenden Proteinen befinden sich zum größten Teil membranassoziierte Proteine, wie Transportproteine und Transporteruntereinheiten oder Enzyme der Atmungs- und Elektronentransportkette. Dieser Typ von Proteinen besitzt in der Regel eine Transmembrandomäne als Teil ihrer Aminosäuresequenz, mit der sie in den unterschiedlichen Membranen der Zelle verankert oder eingebettet sind. Die Detektion solcher Proteinsequenzen ist jedoch, bedingt durch die Extraktions- und Detektionsmethode, mit der die Daten gewonnen wurden, nicht möglich.

Bei Betrachtung der übrigen, nicht erfassten Proteine des Netzwerkes, ist festzustellen, dass circa ein Drittel dieser Proteine beziehungsweise deren Funktionen, mehrfach im metabolischen Netzwerk vorhanden sind (Isoenzyme). Dies hat zur Folge, dass die Aktivität dieser Enzyme beziehungsweise deren Reaktionen für das Netzwerk und die Simulationen nicht relevant sind. Die Existenz solcher Gene kann dadurch erklärt werden, dass es sich um Pseudogene handelt, die niemals transkribiert werden oder deren Transkription nur unter sehr speziellen Bedingungen erfolgt, diese aber nicht bei der Gewinnung der Daten zugegen waren.

Es ist zu beachten, dass trotz der Detektion des Genprodukts der meisten metabolischen Enzyme durch die Auswertung der Proteindaten und deren Verknüpfung mit dem metabolischen Netzwerk, keine direkte Aussage über die Aktivität dieser Enzyme möglich ist. Posttranslationale Modifikationen, welche die Aktivität dieser Enzy-

me stark beeinflussen, können anhand einer solchen Analyse nicht aufgeschlüsselt werden. Zudem enthalten die Proteindaten keine Information über die Aktivität einer spezifischen Reaktion beziehungsweise ob über den beteiligten Stoffwechselweg ein Fluss *in vivo* läuft.

Nichtsdestotrotz ist es ein erster Schritt, um genomskalige Netzwerke zu verifizieren und um deren Aussagekraft bezüglich der Simulationen höher einzuschätzen. Ferner erlauben die Schlüsse aus solch einem systematischen Ansatz der Netzwerkanalyse mit der Verknüpfung von experimentell gewonnenen Daten, ein konkreteres Bild über den Zustand des Zellstoffwechsels *in vivo*.

Modellierung verschiedener Biokraftstoffsynthesen in Cyanobakterien

Vorwort

Das 5. Kapitel dieser Arbeit befasst sich mit einer biotechnologischen Anwendung des metabolischen Modells aus dem Bereich des *Metabolic Engineering*. Konkret wird die Synthese von neun unterschiedlichen Biokraftstoffen analysiert, deren jeweiliger Syntheseweg in das metabolische Netzwerk von *Synechocystis* sp. PCC 6803 integriert wurde. Die entsprechenden metabolischen Reaktionen wurden zu diesem Zweck aus der Literatur entnommen. Zu den untersuchten Biokraftstoffen gehören Ethanol, Ethylen, Propan, Butanol, Isopren, Pentadecan, Heptadecan, Octadecansäure und Octadecanol.

Unter Verwendung der Flussbilanzanalyse wurden die jeweiligen Synthesewege charakterisiert und der zugehörige theoretische maximale Energiegewinn für jeden Biokraftstoff berechnet. Zudem werden auftretende synergetische Wechselwirkungen zwischen dem Biokraftstoff und dem Zellwachstum analysiert. Abschließend wird eine Methode zur Identifizierung von Zielgenen (*Gene Targets*) beziehungsweise Schlüsselreaktionen, die sich für eine genetische Manipulation in Form eines Gen-Knockouts oder einer Genüberexpression zur Steigerung der Produktausbeute eignen, am Beispiel von Ethylen kurz vorgestellt.

5.1. Material und Methoden

5.1.1. Integration von Biokraftstoffsynthesen in das metabolische Netzwerk von *Synechocystis* sp. PCC 6803

Für die Analyse der Biokraftstoffsynthesen wurde die Netzwerkrekonstruktion von *Synechocystis* sp. PCC 6803 mit spezifischen Genen und deren Reaktionen zur Erzeugung von neun unterschiedlichen Biokraftstoffen erweitert. Zu den untersuchten Biokraftstoffen gehören Ethanol, Ethylen, Propan, Butanol, Isopren, Pentadecan, Heptadecan, Octadecansäure und Octadecanol. Die zugehörigen Reaktionen wurden aus der Literatur entnommen und das metabolische Netzwerk von *Synechocystis* um diese Reaktionen erweitert. Zusätzlich wurde für jeden Biokraftstoff eine Transportfunktion integriert, die das jeweilige Produkt aus der Zelle exportieren kann. Die zusätzlichen Synthesereaktionen der einzelnen Biokraftstoffe sind in Tabelle 5.1 aufgeführt.

In Abbildung 5.1 ist zudem gezeigt, in welchem Bereich des Stoffwechsels die jeweiligen Synthesereaktionen liegen und welche direkte Vorläufermetaboliten aus dem Zellstoffwechsel benötigt werden. Da bei der Synthese von Ethylen Guanidin entsteht, welches nicht weiter über das Netzwerk verstoffwechselt werden kann, wurde für diesen Metabolit ein Ausfluss aus dem System zusätzlich in das Netzwerk eingefügt.

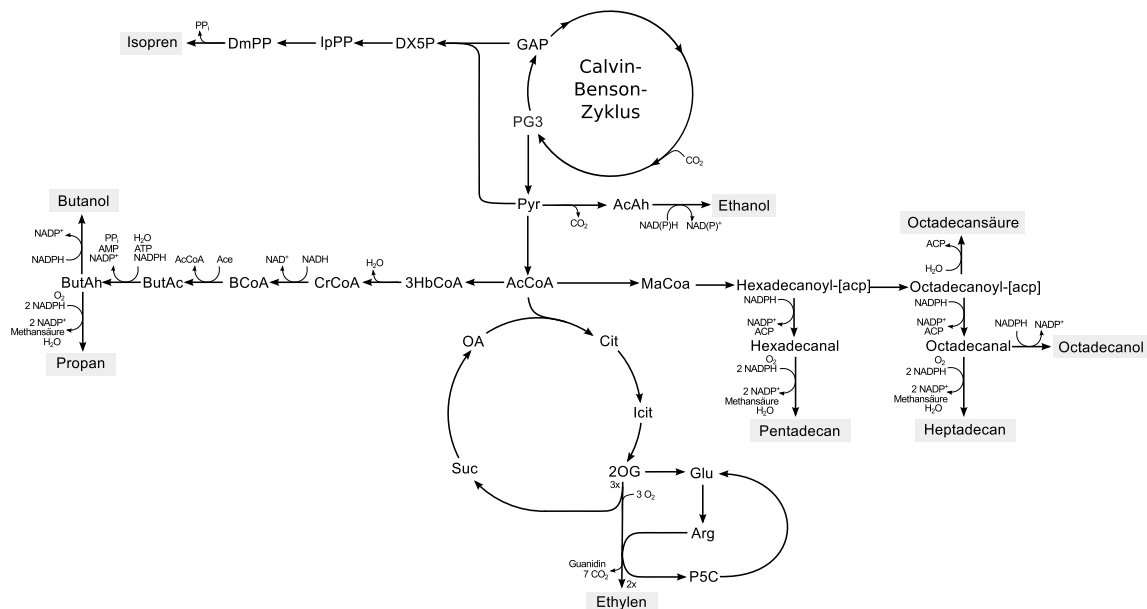


Abbildung 5.1.: Lage der Synthesereaktionen der neun Biokraftstoffe im Zentralstoffwechsel von *Synechocystis* sp. PCC 6803. Die eingefügten Reaktionen sind dabei im Detail mit den jeweiligen Cofaktoren aufgeführt, während die restlichen Reaktionen des Stoffwechsels vereinfacht wurden. Die Erklärung der Abkürzungen für die verwendeten Metaboliten sind im Anhang in der Tabelle A.3 aufgeführt.

Tabelle 5.1.: Synthesereaktionen der neun Biokraftstoffe, die in das metabolische Netzwerk von *Synechocystis* sp. PCC 6803 integriert wurden. Falls vorhanden, sind jeweils die zugehörigen Gene, Enzyme und die entsprechende Referenz zu der Literatur angegeben.

Biokraftstoff	Enzym	Reaktion	Gen	Referenz
Ethanol	Pyruvate-Decarboxylase	Pyruvat \Rightarrow Acetaldehyd + CO ₂		
	Alkohol-dehydrogenase	Acetaldehyd + NAD(P)H + H ⁺ \Rightarrow Ethanol + NAD(P) ⁺	slr0090, slr1192	[226]
Ethylen	Ethylen-bildendes Enzym (EFE)	3 2-Oxoglutarat + 3 O ₂ + L-Arginin \Rightarrow 2 Ethylen + 7 CO ₂ + Succinat + Guanidin + 1-Pyrrolin-5-Carboxylat		[91]
Propan, Butanol	3-Hydroxy-Butyryl-CoA-Dehydratase	Hydroxy-Butyryl-CoA \Rightarrow Crotonoyl-CoA + H ₂ O		[82]
	Trans-2-Enoyl-CoA-Reduktase	Crotonoyl-CoA + NADH + H ⁺ \Rightarrow Butyryl-CoA + NAD ⁺		[82]
	Butyryl-CoA-Transferase	Butyryl-CoA + Acetat \Rightarrow Butansäure + Acetyl-CoA		[82]
		Butansäure + NADPH + H ⁺ + ATP \Rightarrow Butyraldehyd + NADP ⁺ + AMP + PP _i + H ₂ O		
		Butyraldehyd + O ₂ + 2 NADPH + 2 H ⁺ \Rightarrow Propan + Methansäure + H ₂ O + 2 NADP ⁺		
Isopren	Isopren-Synthase	Butyraldehyd + NADPH + H ⁺ \Rightarrow Butanol + NADP ⁺		
		Dimethylallylpyrophosphat \Rightarrow Isopren + PP _i		[10]
Heptadecan, Octadecanol, Pentadecan	Acyl-ACP-Reduktase	Hexadecanoyl-[acp] + NADPH + H ⁺ \Rightarrow Hexadecanal + ACP + NADP ⁺	sl10209	[74]
		Octadecanoyl-[acp] + NADPH + H ⁺ \Rightarrow Octadecanal + ACP + NADPH ⁺	sl10209	[74]
	Aldehyd-Decarbonylase	Hexadecanal + O ₂ + 2 NADPH + 2 H ⁺ \Rightarrow Pentadecan + Methansäure + H ₂ O + 2 NADP ⁺	sl10208	[227]
	Aldehyd-Decarbonylase	Octadecanal + O ₂ + 2 NADPH + 2 H ⁺ \Rightarrow Heptadecan + Methansäure + H ₂ O + 2 NADP ⁺	sl10208	[227]
		Octadecanal + NADPH + H ⁺ \Rightarrow Octadecanol + NADP ⁺		
Octadecan-säure	Acyl-ACP-Thioesterase	Octadecanoyl-[acp] + H ₂ O \Rightarrow ACP + Octadecansäure	slr1609	[83]

5.1.2. Parameter und Berechnung der stöchiometrischen Eigenschaften für die Biokraftstoffsynthesen

Für die Berechnung der stöchiometrischen Eigenschaften, der Trade-off-Bestimmung und des synergetischen Effektes der einzelnen Biokraftstoffe wurde die RuBisCO-Carboxylierungsreaktion von der Oxygenierungsreaktion entkoppelt, so dass diese separat und unabhängig voneinander ablaufen können.

Als Ausgangsszenario wurden ähnliche Parameter für das metabolische Netzwerk festgelegt, wie die in Kapitel 2.1.8 beschriebenen (unter anderem eine Verdopplungszeit von 24 h und keine Aufnahme von komplexeren Kohlenstoffverbindungen), allerdings mit einigen Ausnahmen: *Synechocystis* ist in der Lage, Hydrogencarbonat über verschiedene Transporter aufzunehmen, unter anderem über einen ATP-abhängigen Transport und über einen Na⁺-abhängigen Transport. Im folgenden wird davon ausgegangen, dass Hydrogencarbonat über den Na⁺-abhängigen Transport aufgenommen wird und CO₂, welches im Laufe der unterschiedlichen Synthesewege frei wird, über die Zellmembran ins Medium diffundiert. Ein alternatives Szenario ist ein ATP-abhängiger Transport. Hierbei wird das freigewordene CO₂ nicht aus dem System exportiert. Stattdessen findet ein internes Recycling zu Hydrogencarbonat unter Verwendung des NADPH-Dehydrogenase-Komplexes (NDH-1) statt. Außerdem wurden alle zusätzlichen energieverbrauchenden Reaktionen, wie die Reaktion für den basalen ATP-Verbrauch (*Maintenance*), die Bildung von reaktiven Sauerstoffspezies, die

Tabelle 5.2.: Eingangsparameter für die Simulationen des metabolischen Netzwerkes zur Biokraftstoffsynthese. Aufgeführt sind die festgelegten Flussraten [mmol gDW⁻¹ h⁻¹] für die entsprechenden Reaktionen im Netzwerk. Als zu maximierende Zielfunktion wird entweder die jeweilige Biokraftstoffsynthesereaktion definiert oder die Biomassebildung (R147).

Reaktion	Reaktion	Wert
Lichteingang	R443	15,57
Diverse Kohlenstoff-transporter	R676	0
	R677	0
	R678	0
	R679	0
	R680	0
	R681	0
	R682	0
	R685	0
	R686	0
	R687	0
	R712	0
	R713	0
Zielfunktion	Biokraftstoff bzw. Biomasse	max.

Mehler-ähnliche-Reaktion oder die Respiration, nicht festgelegt. Die gesetzten Reaktionsbegrenzungen des metabolischen Netzwerkes für die Simulationen zu den Biokraftstoffsynthesen sind in der Tabelle 5.2 aufgeführt.

Die optimale Flussverteilung wurde anschließend mithilfe der FBA (siehe Kapitel 2.1.2) berechnet und lieferte die jeweilige maximale, theoretische Biokraftstoffausbeute und die stöchiometrische Anzahl der Photonen, von Sauerstoff, NADPH, ATP sowie der Kohlenstoffatome in Millimol, die über die Reaktion der RuBisCO fixiert werden, um ein Millimol des spezifischen Biokraftstoffes zu synthetisieren. Zudem ist die Gesamtanzahl der aktiven Reaktionen im Netzwerk pro Biokraftstoff gezeigt. Die jeweiligen Heizwerte der einzelnen Biokraftstoffe bei Raumtemperatur (H_i) wurden aus dem NIST Chemie-Webbuch [228] entnommen. Mithilfe der maximalen Syntheserate des Biokraftstoffes und des zugehörigen Heizwertes wurde der maximale theoretische Ertrag durch Multiplikation beider Werte errechnet.

5.1.3. Trade-off und Variabilitätsanalyse der Biokraftstoffsynthesen

Um den Energie Trade-off bei den verschiedenen prozentualen Biomassessynthesanteilen zu berechnen, wurde die Synthese der Biomasse zwischen 0 % und 100 % der maximalen Biomassessynthese festgelegt. Startend mit der maximalen Biomassessynthese wurde diese bei jedem Rechenschritt schrittweise verringert. Anschließend wurde das System auf die Synthese des gewünschten Biokraftstoffes optimiert, so dass alle verbleibenden Ressourcen auf dessen Produktion gerichtet wurden. Dabei wurde für jede einzelne Reaktion die Flussvariabilität berechnet, die möglich ist, um die Maximierung der Zielfunktion zu gewährleisten. Ausgewählt wurden abschließend die Reaktionen, bei denen entweder der absolute Flusswert, in diesem Fall der Durchschnitt der Variabilität, mit ansteigender Biokraftstoffsynthese um mindestens 5 % über den gesamten Verlauf ansteigt beziehungsweise abfällt und bei 100 %iger Biokraftstoffsynthese mindestens $0,001 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ absolut beträgt.

5.1.4. Synergetische Effekte der Biokraftstoffsynthesen

Neben der alleinigen Synthese des Biokraftstoffes wurde auch die Ausbeute analysiert, wenn die Bildung von Biomasse und die Synthese des Biokraftstoffes gleichzeitig in der Zelle ablaufen. Der bei den unterschiedlichen Prozentschritten der Biokraftstoffsynthese (siehe Kapitel 5.1.3) resultierende Synthesefluss wurde mit dem maximalen Synthesefluss ohne Biomassebildung bei der gleichen Lichtintensität verglichen und der prozentuale Anteil errechnet. Anhand dieses Vergleiches lassen sich synergetische Effekte des jeweiligen Biokraftstoffes ableiten. Ein synergetischer Effekt soll in diesem Zusammenhang bedeuten, dass der restliche prozentuale Anteil an Ressourcen, der nicht in die Synthese von Biomasse eingeht, einen höheren prozentualen Synthesefluss des Biokraftstoffes erlaubt, als zu erwarten wäre.

Für die Berechnung des optimalen Verhältnisses mit den maximalen synergetischen Effekten zwischen Biomasseproduktion und Kraftstoffsynthese wurde die folgende Funktion maximiert:

$$\max \frac{v_{fuel}^{max}}{v_{bio}^{max}} v_{bio} + v_{fuel},$$

wobei v_{fuel} und v_{bio} die Syntheserate des jeweiligen Biokraftstoffs bzw. der Biomasse darstellen. Somit wird gleichzeitig die Biomassebildung als auch die Biokraftstoffsynthese optimiert.

5.2. Ergebnisse und Diskussion

5.2.1. Stöchiometrische Modellierung von verschiedenen Biokraftstoffsynthesen

Die Eignung eines potenziellen Biokraftstoffes wird neben der Toxizität und anderen Faktoren hauptsächlich durch die maximale stöchiometrische Ausbeute im Verhältnis zu dem jeweiligen Heizwert beschrieben. Es konnte bereits gezeigt werden, dass die maximale stöchiometrische Ausbeute anhand von genomskaligen Stoffwechselrekonstruktionen, zum Beispiel von Cyanobakterien [145], berechnet werden kann. Mithilfe dieser Stoffwechselrekonstruktionen können relevante Faktoren für die Biosynthese, wie stöchiometrisch optimale Synthesewege, identifiziert werden. Die Flussbilanzanalyse (*Flux Balance Analysis*) (FBA) ermöglicht es, durch die Anwendung der linearen Programmierung auf biologische Probleme, optimale Flussverteilungen anhand von Stöchiometrie und anderen Kriterien, die den Lösungsraum einschränken, zu berechnen [229].

Unter Verwendung der Rekonstruktion und dem stöchiometrischen Modell des Stoffwechsels von *Synechocystis* sp. PCC 6803 (siehe Kapitel 2) und der FBA konnten für neun Biokraftstoffe mehrere stöchiometrische Eigenschaften berechnet werden (siehe Tabelle 5.3). Der Fokus lag dabei auf den Biokraftstoffen Ethanol, Ethylen, Propan, Butanol, Isopren, Pentadecan, Heptadecan, Octadecansäure und Octadecanol.

Die untersuchten Produkte unterscheiden sich in Bezug auf die Anzahl an Kohlenstoffatomen und deren Heizwert (H_i). Zuerst wurden die metabolischen Anforderungen berechnet, um ein Molekül des jeweiligen Biokraftstoffes zu synthetisieren, ohne dabei ein Zellwachstum zu berücksichtigen. Von besonderem Interesse sind hierbei die Anzahl der involvierten Reaktionen, die Photonenanzahl, die Freisetzung von molekularem Sauerstoff, der Bedarf an NADPH und ATP sowie der Fluss durch die RuBisCO-Carboxylase-Reaktion.

Tabelle 5.3.: Stöchiometrische Eigenschaften der neun analysierten Biokraftstoffe jeweils pro Molekül Biokraftstoff. Aufgeführt sind dabei die Anzahl der Kohlenstoffatome (C-Atome), der Heizwert (H_i), die Anzahl der Reaktionen im Netzwerk, die bei 100-prozentiger Biokraftstoffsynthese einen Fluss tragen, der Bedarf an Photonen, NADPH und ATP, die Anzahl der produzierten Sauerstoffmoleküle (O_2) und der fixierten CO_2 -Moleküle durch RuBisCO sowie die maximal mögliche Biokraftstoffsyntheserate und der jeweilige Energiegewinn.

Bio- kraftstoff	C- Atome	H_i [kJ mol ⁻¹]	Reaktionen	Photonen	NADPH	ATP	ATP: NADPH	O_2	RuBisCO	max Synthesefluss [mol gDW ⁻¹ h ⁻¹]	Energie- gewinn [J gDW ⁻¹ h ⁻¹]
Ethanol	2	1370	51	24	6	5	0,83	3	2	6,49E-04	888,79
Ethylen	2	1410	82	61	9,5	19,5	2,05	6	6	2,55E-04	359,90
Propan	3	2220	66	52	13	12	0,92	5,5	4	2,99E-04	664,72
Butanol	4	2670	62	48	12	12	1,00	6	4	3,24E-04	866,08
Isopren	5	3200	66	56	13	17	1,31	7	6	2,78E-04	889,71
Pentadecan	15	10000	92	196	49	55	1,12	23,5	16	7,21E-05	720,83
Heptadecan	17	11400	96	220	55	62	1,13	26,5	18	7,49E-05	853,36
Octadecansäure	18	11300	92	208	52	62	1,19	26	18	7,08E-05	799,73
Octadecanol	18	11820	93	216	54	62	1,15	27	18	7,94E-05	938,97

Der Wert der RuBisCO-Carboxylase-Reaktion kann sich von der Anzahl der Kohlenstoffmoleküle pro Einheit des Biokraftstoffes unterscheiden, da oft CO₂ oder andere Nebenprodukte während der Biosynthese des Produktes frei werden und der Kohlenstoff neu fixiert wird. Um die theoretische Ausbeute der untersuchten Biokraftstoffe direkt zu vergleichen, ist zudem die maximale Syntheserate und der theoretische Energiegewinn in der Tabelle 5.3 aufgeführt.

Für einen definierten Lichteinfall besitzt Ethanol die höchste Syntheserate, gefolgt von Butanol, Propan und Isopren. Anschließend folgen Pentadecan, Octadecansäure, Octadecanol und Heptadecan. Es ist zu erkennen, dass eine klare Korrelation zwischen der Länge des Syntheseweges, der Photonen-Aufnahme, des NADPH- und ATP-Bedarfs, der O₂-Freisetzung und der maximalen Syntheserate existiert. Allerdings zeigt die Länge des Syntheseweges keine Korrelation zu dem Verhältnis von ATP zu NADPH. Bei einer genaueren Betrachtung des Verhältnis zwischen ATP und NADPH ist zu erkennen, dass dieses zwischen 0,86 und 2,05 variieren kann. Die meisten der betrachteten Biokraftstoffe weisen ein ähnliches Verhältnis auf (zwischen 0,86 und 1,31) mit Ausnahme von Ethylen, bei dessen Synthese mehr als doppelt soviel ATP wie NADPH benötigt wird. Das optimale Verhältnis, welches durch den linearen Elektronenfluss an der Thylakoidmembran vorgesehen ist, liegt bei Simulationen im Modell bei 1,29. Bei den berechneten Werten aus Tabelle 5.3, welche dem maximalen stöchiometrischen Potential entsprechen, ist zu beachten, dass diese in Abwesenheit von konkurrierenden Stoffwechselwegen, eines basalen Energieverbrauchs (*Maintenance*) oder energieverschwendenden Zyklen erzielt wurden, welche natürlich diese optimalen Werte beeinflussen können.

Der theoretische Energiegewinn ergibt sich aus dem maximalen Synthesefluss und dem jeweiligen Heizwert des Biokraftstoffes. Aufgrund des hohen Heizwertes verfügt Octadecanol über den höchsten Wert. Darauf folgen Isopren, Ethanol, Butanol, Heptadecan, Octadecansäure, Pentadecan, Propan und Ethylen. Es ist auszumachen, dass die Alkane im Vergleich zu den Alkoholen, einen niedrigeren Energiegewinn aufweisen. Dieser Sachverhalt korreliert mit der Anzahl der freigesetzten Kohlenstoffatome während der Synthese. Diese Tatsache lässt sich dadurch erklären, dass bei der Alkansynthese die Aldehyd-Decarboxylase (sll0208) [227] Methansäure freisetzt (siehe Kapitel 5.1.1), wodurch vorher fixierter Kohlenstoff verloren geht und dadurch der maximale Synthesefluss verringert wird.

5.2.2. Trade-off zwischen Wachstum und Biokraftstoffsynthese

Für die Herstellung von Biokraftstoffen in großem Maßstab ist es wichtig, dass die produzierenden Zellen auch weiterhin lebensfähig bleiben. Eine vollkommene Biokraftstoffsynthese, bei der jegliche Ressourcen in die Kraftstoffsynthese fließen, wird *in vivo* in der Zelle nicht durchführbar sein. Für ein Überleben der Zelle ist es notwendig, dass die zellulären Funktionen, wie unter anderem ein minimales Zellwachstum und ein zellulärer basaler Energieverbrauch (*Maintenance*), erhalten bleiben. Die Aufrechterhaltung dieser Funktionen kostet Energie und reduziert zwangsweise den Anteil,

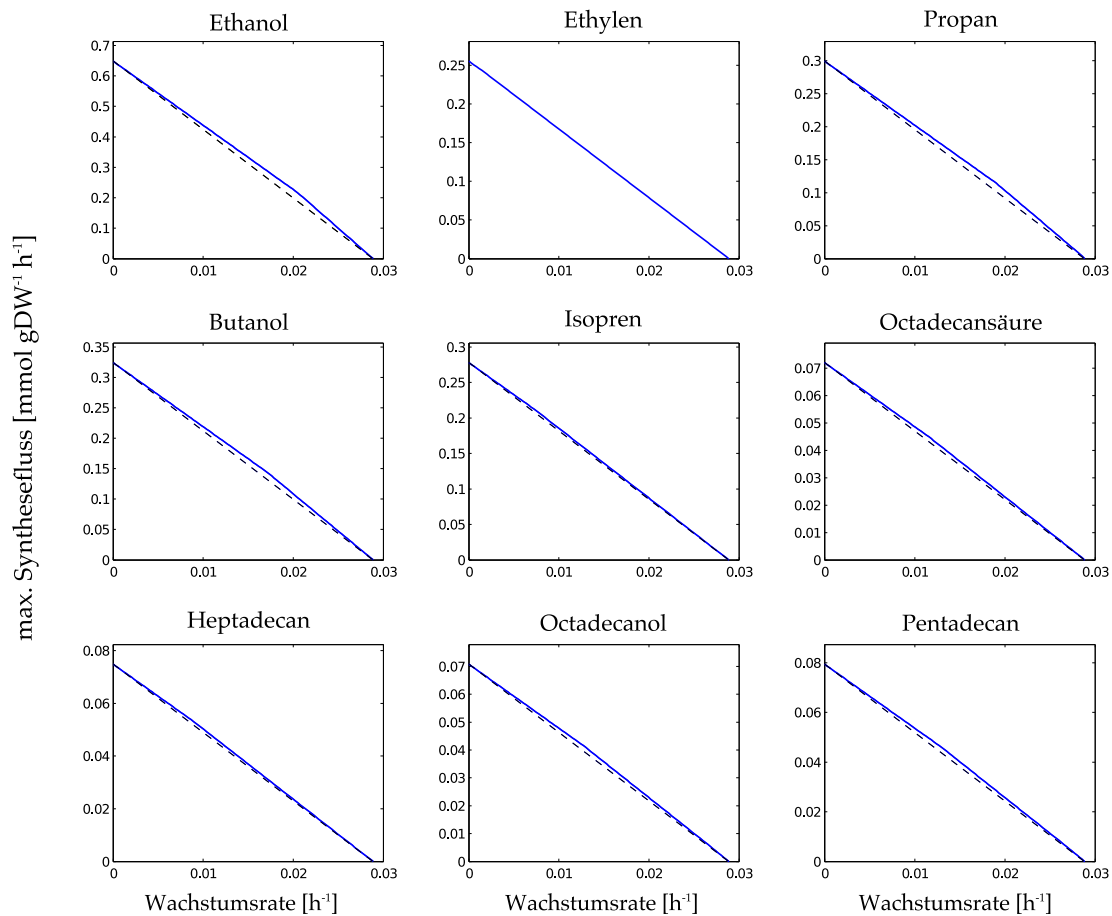


Abbildung 5.2.: Trade-off zwischen Wachstum und Biokraftstoffsynthese. Gezeigt ist die maximale Syntheserate der neun untersuchten Biokraftstoffe in Abhängigkeit zur Wachstumsrate beziehungsweise Biomassensynthese. Der lineare Verlauf zwischen maximaler Biokraftstoffsynthese und Biomassensynthese ist als gestrichelte Hilfslinie eingezeichnet.

der in die Synthese des gewünschten Kraftstoffproduktes fließen kann. Ein reduzierter Ertrag des Zielprodukts ist die Folge.

Beginnend mit der optimalen Flussverteilung des Wildtyp-Wachstums, kann der Fluss schrittweise zur Synthese des gewünschten Biokraftstoffes gelenkt werden. Die maximale Wachstumsrate wird reduziert, bis alle zellulären Ressourcen in die Synthese des Biokraftstoffes fließen. Dabei ist es von Interesse, ob ein Optimum existiert, bei dem ausreichend Biomasse gebildet werden kann, bei gleichzeitiger Maximierung der Produktion des Biokraftstoffes. Bei Miteinbeziehung des Zellwachstums in die stöchiometrische Analyse ist zu erkennen, dass generell ein Trade-off zwischen der Synthese der gewünschten Endprodukte und der Bildung von Biomasse besteht (dargestellt in Abbildung 5.2).

Biokraftstoffe, bei denen wenig oder gar kein Kohlenstoff während der Synthese verloren geht, zeigen ein gutes Energie -Trade-off- Profil. Andere Biokraftstoffe, die einen

hohen Verlust von Kohlenstoffatomen während des Syntheseprozesses, im Vergleich zu deren ursprünglicher Anzahl an Kohlenstoffatomen, aufweisen, wie zum Beispiel Ethylen, besitzen ein sehr schlechtes Profil. Wenn Kohlenstoff verloren geht, ist der Trade-off besser, wenn die Anzahl der Kohlenstoffatome im jeweiligen Biokraftstoff beziehungsweise das Verhältnis zwischen diesen und dem freigesetzten Kohlenstoff hoch ist. Aber auch der hohe Bedarf an ATP und NADPH spielt eine wichtige Rolle, wodurch sich der maximale Synthesefluss stark verringert und somit auch der maximale theoretische Energiegewinn niedrig bleibt.

Bei einer genauen Betrachtung der Trade-off-Profile (Abbildung 5.2) ist zu erkennen, dass der errechnete Biokraftstoffsynthesefluss (durchgezogene Linie) nicht exakt einem linearen Anstieg von der maximalen Biomassesynthese (Wachstumsrate) gleicht, sondern dieser in den meisten Fällen etwas höher liegt. Der lineare Verlauf (gestrichelte Linie) repräsentiert den zu erwartenden Synthesefluss. Dies bedeutet, dass der prozentuale Anteil der Ressourcen, der nicht in das Wachstum fließt, zur Synthese des Kraftstoffes zum gleichen Anteil der maximalen Synthese zur Verfügung steht. Bei einer 100-prozentigen Biomassesynthese würde dies beispielsweise 0 % der maximalen Biokraftstoffsynthese bedeuten, bei 75 % Biomassesynthese dürften 25 % der maximalen Kraftstoffsynthese erwartet werden. Die Simulationen ergeben jedoch, dass eine höhere Biokraftstoffsynthese möglich ist, als zu erwarten wäre (siehe Abbildung 5.2). Ausschlaggebend für diesen Sachverhalt sind synergetische Effekte, die zwischen der Synthese des Biokraftstoffes und der Biomasse auftreten. Diese Effekte variieren in ihrem Ausmaß für die unterschiedlichen Biokraftstoffe bei unterschiedlichen prozentualen Syntheseraten. So sind zum Beispiel bei der Ethanol synthese größere synergetische Effekte zu erkennen als bei der Ethylen- oder Isoprensynthese. Eine genauere Analyse dieser Effekte wird im nächsten Abschnitt beschrieben.

5.2.3. Synergetische Effekte bei der Biokraftstoffsynthese

Bei der Analyse verschiedener Biokraftstoffe gibt es zweifelsohne einen Unterschied, je nachdem ob die Synthese eines Kraftstoffes für sich separat betrachtet wird (wie in Tabelle 5.3) oder sich diese in einem bereits etablierten biologischen System abspielt. Bei Letzterem ist es von großem Interesse, welchen Einfluss die Synthese des Biokraftstoffes auf das System hat und welche Wechselwirkungen wiederum dabei existieren. Insbesondere können Nebenprodukte, die während der Synthese des Biokraftstoffes auftreten, bei einer simultanen Betrachtung des Zellwachstums zurück in den Stoffwechsel kanalisiert und dort für das Wachstum recycelt werden. Bei der Betrachtung der Synthese des Biokraftstoffes ohne das Zellwachstum stellen diese Nebenprodukte, wie zum Beispiel Methansäure, die während der Alkan-Synthese frei wird, eher eine metabolische Belastung dar.

Für eine Untersuchung der synergetischen Wechselwirkung zwischen der Bildung von Biomasse und Biosynthese des Biokraftstoffes wurde der Synthesefluss unter alleiniger Bildung des Biokraftstoffes mit dem Fluss, wenn das Zellwachstum auf einen prozentualen Anteil des maximalen Wachstums festgelegt wird, verglichen. Der pro-

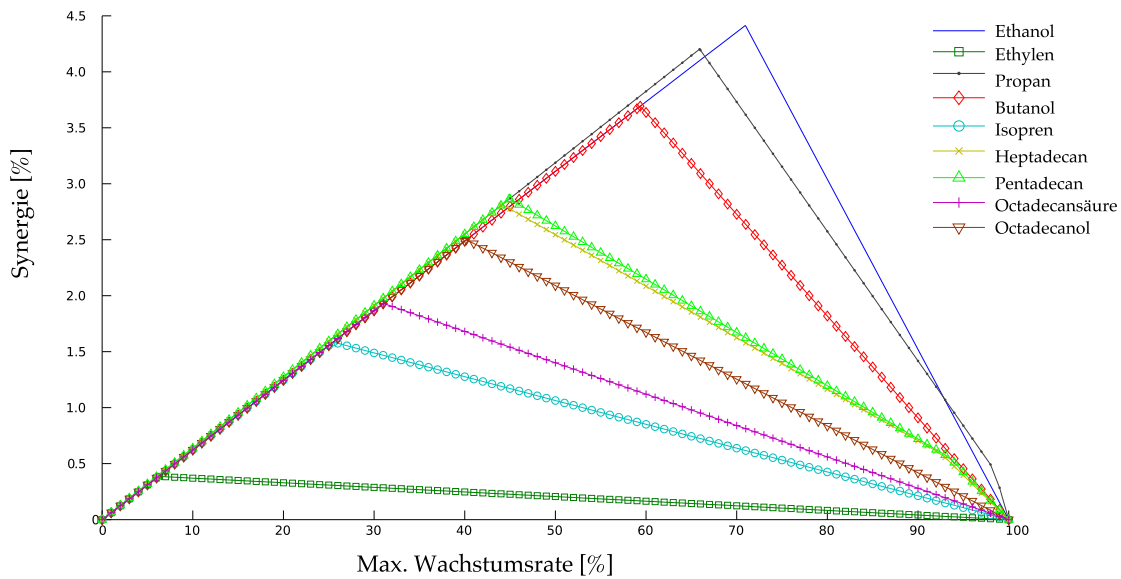


Abbildung 5.3.: Analyse des Synergieeffekts der unterschiedlichen Biokraftstoffe. Für die neun Biokraftstoffe wurde jeweils die Biomassesynthese prozentual erhöht und simultan auf die Synthese des Biokraftstoffes optimiert. Die Synergie entspricht hierbei dem prozentualen Anteil der Biokraftstoffsynthese, der über dem zu erwartenden Anteil liegt.

zentuale Anteil, der über dem restlichen, zu erwartenden Anteil des maximalen Flusses liegt, wird als Synergie ausgedrückt.

Um Rückschlüsse zu erhalten, bei welchem Verhältnis zwischen Biomasse und Kraftstoffsynthese die größte Synergie herrscht, wurden beide Synthesen, also die von Biomasse und Biokraftstoff, gleichzeitig optimiert und für jeden Biokraftstoff das Optimum berechnet. Dieses ermöglicht einen direkten Vergleich, welche Biokraftstoffe die besten metabolischen und energetischen Vorteile im Idealfall und in einem der Realität beziehungsweise des *in vivo*-Zustands nachempfundenen Fall besitzen. In Abbildung 5.3 sind die Synergiewerte für alle neun Biokraftstoffe in Abhängigkeit zum Anteil der Biomassesynthese aufgetragen.

In Abbildung 5.3 ist zu erkennen, dass Ethanol, Propan und Butanol das Maximum der Synergie bei einer prozentual höheren Biomasseproduktion (60-75 %) erreichen, während zum Beispiel die Alkane wie Penta- und Hexadecan, die über eine Aldehyd-Decarbonylase synthetisiert werden, eher im Bereich um 40-50 % ihr Maximum besitzen. Für die restlichen Kraftstoffe liegt das Maximum bei noch niedrigeren prozentualen Biomasseraten und auch der maximale Synergiewert an sich ist bei diesen geringer.

Ein entscheidender Faktor, der in diese Analyse mit einfließt, ist das Verhältnis zwischen ATP und NADPH, welches für die Synthese des jeweiligen Kraftstoffes benötigt wird. Ethanol, Propan und Butanol benötigen eher ein niedriges Verhältnis von ≤ 1 , wohingegen die restlichen Kraftstoffe mehr Moleküle ATP als NADPH benötigen. Im metabolischen Modell liefern die Photosysteme inklusive der Elektronentransportkette im optimalen Fall ein Verhältnis von 1,29:1. Dies hat zur Folge, dass

bei einer 100-prozentigen Biokraftstoffsynthese, die Elektronentransportkette nicht optimal genutzt wird. Wenn jedoch zu einem bestimmten Anteil Biomasse simultan synthetisiert wird - die Synthese von Biomasse besitzt ein ATP-NADPH-Verhältnis von 2,13:1 - gewährleistet dies eine optimale Ausnutzung der Elektronentransportkette, wodurch der prozentuale Anteil der Biokraftstoffausbeute steigt und dementsprechend ein synergetischer Effekt im Modell erzielt wird.

Neben dem Verhältnis zwischen ATP zu NADPH spielen auch Nebenprodukte, die während der Biokraftstoffsynthese entstehen und für das Wachstums genutzt werden können, eine Rolle. Bei der Synthese der Alkane, also Propan, Pentadecan und Heptadecan, entsteht durch die Reaktion der Aldehyd-Decarboxylase (sll0208) [227] Methansäure. Bei einer 100-prozentigen Biokraftstoffsynthese, kann diese Methansäure nicht weiter im Stoffwechsel genutzt werden und wird über die Reaktion der Formatdehydrogenase (sll1359) [119] zu CO₂ oxidiert. Treten andererseits Biokraftstoff- und Biomassesynthese gleichzeitig auf, kann die während der Kraftstoffsynthese freigesetzte Methansäure durch die Reaktion der Phosphoribosylglycinamid-Transformylase 2 (EC 2.1.2.-) (*purT*, slr0861) für die Purinsynthese und daran anschließende Stoffwechselwege weiterverwendet werden, um effektiver Biomasse zu produzieren.

Bei der Analyse der Synergie ist zu beachten, dass für die Simulation alle zusätzlichen energieverbrauchenden Reaktionen, wie der basale ATP-Verbrauch (*Maintenance*), die Bildung von reaktiven Sauerstoffspezies, die Mehler-ähnliche-Reaktion oder die Zellrespiration, nicht festgelegt wurden und somit gleich null waren. Werden jedoch diese Prozesse bei der Simulation berücksichtigt und die Parameter für die jeweiligen Reaktionen im Modell festgelegt, sind die synergetischen Effekte aller Biokraftstoffe nur noch sehr gering.

5.2.4. Variabilitätsanalyse zur Identifikation von Schlüsselreaktionen bei der Biokraftstoffsynthese

Neben der Betrachtung der einzelnen Biokraftstoffe auf Ebene des gesamten Netzwerkes (Systems) und dem Zusammenspiel zwischen deren Synthese und dem Zellwachstum, ist das Verhalten von einzelnen Reaktionen des metabolischen Netzwerkes von großem Interesse. Für eine gesteigerte Kraftstoffsynthese birgt die genetische Manipulation des Organismus ein großes Potenzial. Dabei können Genüberexpression, Gen-Stillegung (*Gene Silencing*) oder ganze Gen-Knockouts von metabolischen Enzymen, durch das Umleiten des metabolischen Fluss hin zum gewünschten Produkt, helfen, den Ertrag des Biokraftstoffes zu erhöhen. Zur Identifikation solcher Zielgene beziehungsweise Schlüsselreaktionen (*Gene Targets*), deren Manipulation Auswirkungen auf die Kraftstoffsynthese haben, gewährt eine erneute Betrachtung der Trade-off-Analyse zwischen der Biomasse- und der Biokraftstoffsynthese ein detaillierteres Bild.

Bei der Optimierung von bestimmten Zielfunktionen im Netzwerk birgt jede Reaktion eine gewisse Flussvariabilität in sich, das heißt es existieren eine obere und eine untere Flussgrenze, die bei einer Optimierung der metabolischen Flüsse das Maximum der Zielfunktion gewähren. Mit einer Analyse dieser Variabilität erhält man

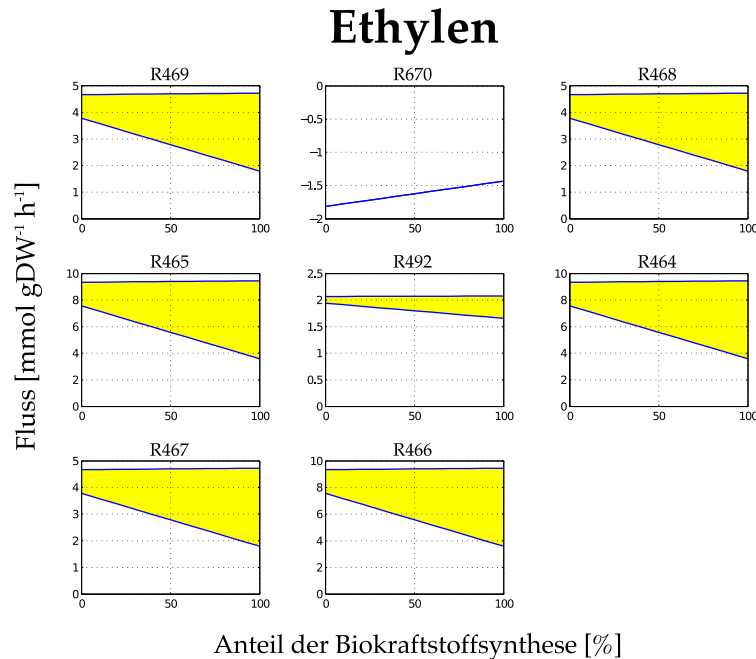


Abbildung 5.4.: Variabilitätsanalyse von Reaktionen mit einer abfallenden Flussaktivität für die Synthese von Ethylen. Gezeigt sind Reaktionen, deren Mittelwerte der Flussvariabilität mit prozentual zunehmender Biokraftstoffsynthese abnehmen und bei maximaler Biokraftstoffsynthese noch einen Fluss tragen. Die Beschreibungen der Reaktionsnummern beziehungsweise welche Reaktionen abgebildet sind, sind im Anhang in der Tabelle A.4 aufgeführt.

Hinweise, welche Reaktionen im Netzwerk essentiell sind und welche Reaktionen verändert werden können, um die Biokraftstoffsynthese zu erhöhen.

Für das metabolische Netzwerkmodell von *Synechocystis* sp. PCC 6803 wurde die Variabilitätsanalyse beim Übergang von der Biomassesynthese hin zur Biokraftstoffsynthese für alle neun analysierten Biokraftstoffe durchgeführt. Im Folgenden werden stellvertretend für alle neun Biokraftstoffe die Ergebnisse am Beispiel der Synthese von Ethylene aufgeführt und diskutiert.

Alle Reaktionen, die mit prozentual ansteigender Ethylensynthese an Aktivität verlieren beziehungsweise deren Mittelwerte der Flussvariabilität sich verringern und die bei maximaler Biokraftstoffsynthese noch einen Fluss tragen, sind in Abbildung 5.4 gezeigt. Insgesamt sind dies für die Synthese von Ethylen acht Reaktionen, bei denen bei maximaler Biokraftstoffsynthese noch ein Fluss vorhanden ist. Davon sind sechs Reaktionen (R464-R469) mit dem Cytochrom-*b₆f*-Komplex assoziiert, die Reaktion R492 steht für die Reaktion der ATPase und die Reaktion R670 repräsentiert die Sauerstoffaufnahme im System. Diese Reaktionen beziehungsweise deren zugehörigen Gene stellen potenzielle Kandidaten für eine genetische Modifikation am Organismus dar, zum Beispiel in Form von Gen-Knockdown-Manipulationen, um den Fluss in Richtung des gewünschten Produktes zu leiten und dessen Ausbeute zu steigern.

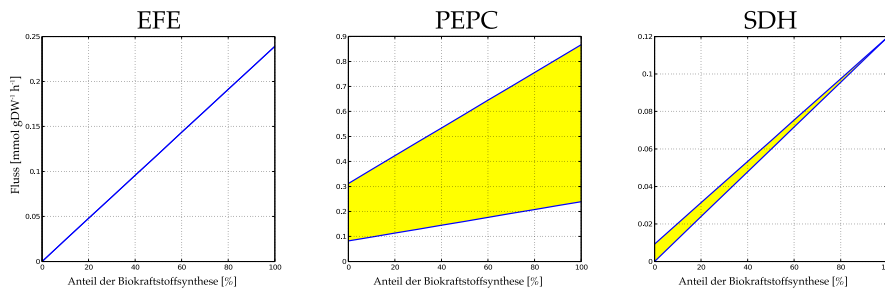


Abbildung 5.5.: Variabilitätsanalyse des Ethylen-bildenden Enzyms (EFE), der PEP-Carboxylase (PEPC) und der Succinat-Dehydrogenase (SDH) für die Synthese von Ethylen. Bei allen drei Reaktionen steigt der Mittelwert der Flussvariabilität mit prozentual zunehmender Ethylensynthese.

Die Anzahl der Reaktionen, die mit ansteigender Ethylensynthese an Aktivität gewinnen, ist mit 89 um ein vielfaches höher als die Reaktionen, die an Aktivität verlieren und bei maximaler Biokraftstoffsynthese noch einen Fluss tragen. Darunter sind vor allem Reaktionen, die den metabolischen Fluss in Richtung des Produktes führen und Reaktionen, die Nebenprodukte, die bei einer erhöhten Biokraftstoffsynthese auftreten, weiter im Stoffwechsel verarbeiten.

Bei einer genaueren Betrachtung der Reaktionen mit einem ansteigenden Fluss ist es möglich, gezielt Schlüsselreaktionen zu identifizieren, die für eine gesteigerte Produktsynthese essentiell sind, beziehungsweise einen höheren Fluss tragen müssten, als bei der Wildtyp-Flussverteilung (100 % Biomassessynthese). Für das Beispiel Ethylen sind hier kurz drei Reaktionen ausgewählt und dargestellt (siehe Abbildung 5.5). Zum einen die Biokraftstoff-produzierende Reaktion an sich, im Fall von Ethylen das Ethylen-bildende Enzym (EC 1.14.11.34) (EFE), aber auch unter anderem die Reaktion der PEP-Carboxylase (EC 4.1.1.31) (PEPC) und der Succinat-Dehydrogenase (EC 1.3.99.1) (SDH)

Für das EFE ergibt sich aus den Simulationen selbstredend ein linearer Zusammenhang zwischen Aktivität und Biokraftstoffsynthese. Je höher die Flussrate durch diese Reaktion verläuft, desto mehr Ethylen wird produziert. Da dies die einzig Ethylen-produzierende Reaktion im Netzwerk symbolisiert, ergibt sich auch keine Variabilität innerhalb des Flusses dieser Reaktion.

Die PEPC carboxyliert Phosphoenolpyruvat zu Oxalacetat, welches ein essentielles Ausgangsprodukt für den TCA-Zyklus in vielen Organismen darstellt. Ohne die PEPC ist kein Fluss über den TCA-Zyklus bei *Synechocystis* möglich und wichtige Komponenten des Stoffwechsels, wie Aminosäuren oder Cofaktoren, könnten nicht synthetisiert werden. Für das Ethylen-bildende Enzym sind 2-Oxoglutarat und Arginin die Hauptsubstrate. Die Synthese von Arginin verläuft von 2-Oxoglutarat über Glutamat zu Arginin. Somit stellt 2-Oxoglutarat den wichtigsten Ausgangsmetaboliten für die Synthese von Ethylen im Stoffwechsel dar. Demzufolge ist eine erhöhte Aktivität des TCA-Zyklus inklusive der PEPC notwendig, um den Fluss über das EFE zu erhöhen.

Die Reaktion des EFE produziert nicht nur ausschließlich Ethylen, sondern auch zu einem gewissen Anteil Succinat (siehe Abbildung 5.1 oder Tabelle 5.1). Um dieses

Nebenprodukt wieder zurück in den Stoffwechsel zu kanalisieren, müssen die entsprechenden Reaktionen auch einen erhöhten Fluss führen. Für den Fall von Succinat ist daran unter anderem die SDH beteiligt. Sie katalysiert die Oxidation von Succinat zu Fumarat mit FAD als Elektronenakzeptor und stellt somit den ersten Schritt beim Wiederverwerten von Succinat dar. Am Beispiel der SDH wird deutlich, dass nicht nur eine erhöhte Aktivität der zum Biokraftstoff hinführenden Reaktionen von Bedeutung ist, sondern auch Reaktionen, die obligatorische Nebenprodukte, die bei einer gesteigerten Biokraftstoffsynthese unausweichlich auftreten, weiterverarbeiten, eine essentielle Aufgabe übernehmen.

Beide Enzyme, die PEPC und die SDH repräsentieren potenzielle Ziele für zum Beispiel eine Genüberexpression, damit genügend Ausgangsstoffe für die Ethylensynthese bereitgestellt werden, aber auch vermehrt auftretende Nebenprodukte abgebaut werden können.

5.3. Schlussfolgerungen

Phototrophe Mikroorganismen bergen das große Potenzial als eine Ressource zur Gewinnung und Erzeugung von hochwertigen Produkten und Biokraftstoffen, einfach nur mit der Hilfe von atmosphärischem Kohlenstoffdioxid, Sonnenlicht und einigen Nährstoffen, zu dienen. In dieser Hinsicht haben Cyanobakterien als ein möglicher Ausgangspunkt zur Erzeugung von Biokraftstoffen (der vierten Generation), in jüngster Zeit auf sich aufmerksam gemacht.

Cyanobakterien- und Mikroalgen-Systeme haben viele Vorteile gegenüber den herkömmlichen und traditionellen Nutzpflanzen und ihre Etablierung könnte in naher Zukunft auch aus wirtschaftlicher Sicht möglich sein, wenn die letzten biotechnischen, umwelttechnischen und wirtschaftlichen Hürden überwunden werden. Letztendlich zeigen die Ergebnisse, dass Cyanobakterien das Potenzial aufweisen, als ein nachhaltiger Faktor im System von erneuerbaren Energien aufzutreten, um somit bei der Suche nach alternativen Energiequellen zur Eindämmung der globalen Erwärmung dienlich zu sein.

Für die Etablierung eines Cyanobakteriensystems kann eine theoretische Modellierung zur Eignung und Verbesserung des produzierenden Organismus von großem Interesse sein. Um reale Vorhersagen für die Produktivität einzelner Organismen zu treffen, wurde in dieser Arbeit gezeigt, dass durch die Rekonstruktion metabolischer Netzwerke und deren Analyse unter Nutzung zum Beispiel der Flussbilanzanalyse (*Flux Balance Analysis*), welches ein kraftvolles Werkzeug zur Analyse metabolischer Prozesse und deren Optimierung darstellt, konkrete Aussagen und Vorhersagen getroffen werden können. Zwischen der maximalen theoretischen Ausbeute der flüssigen Brennstoffe oder Chemikalien, die anhand des stöchiometrischen Potenzials berechnet wurden und der experimentell beobachteten Ausbeute, existieren nur geringe Unterschiede [230].

Neben der Vorhersage über die theoretische Ausbeute des gewünschten Biokraftstoffes, ermöglicht die stöchiometrische Modellierung, die Eignung eines Organismus, in Bezug auf Integration in den vorherrschenden Stoffwechsel, zu analysieren (Stamm-Auswahl). Dies beinhaltet zum Beispiel, ob synergetische Effekte zwischen der Produktsynthese und dem Zellwachstum beziehungsweise dem Zellstoffwechsel des Wildtyps existieren. Natürlich sollte dies nicht ausschließlich als ausschlaggebendes Kriterium bei der Wahl für den Organismus und für den Biokraftstoff gesehen werden, jedoch ist der Informationsgehalt einer solchen Analyse nicht unwichtig. Schließlich stellen neu integrierte Stoffwechselwege, die in ein bestehendes System eingebracht werden, eine biologische Belastung für den Organismus dar und je entfernter der gewünschte Zustand des Stoffwechsels zum Wildtyp-Stoffwechsel liegt, um so unwahrscheinlicher ist eine erfolgreiche Integration inklusive einer hohen Ausbeute.

Abgesehen von der Betrachtung des gesamtheitlichen Systems besteht außerdem die Option, eine detaillierte Analyse bis hinunter zur Ebene der einzelnen Enzyme und Reaktionen durchzuführen, um zentrale Punkte und potenzielle Bottlenecks, die bei der Biokraftstoffsynthese relevant sind, zu identifizieren. Unter Verwendung der

Variabilitätsanalyse konnte am Beispiel von Ethylen gezeigt werden, welche spezifischen Reaktionen für eine Kraftstoffsynthese von Bedeutung sind, um diese als Kandidaten oder Ziele für eine mögliche genetische Manipulation auszumachen (siehe Abbildung 5.5). Eine solche Manipulation umfasst für Reaktionen, die einen erhöhten Fluss für die Biokraftstoffsynthese aufweisen, eine Überexpression des entsprechenden Gens oder die Reaktionen, die für eine optimale Produktsynthese eine verringerte Flussrate ausweisen sollten, ein Knockout oder Knockdown des zugehörigen Gens.

Es stellte sich heraus, dass nicht nur Reaktionen, die offensichtlich zum gewünschten Zielprodukt führen oder in direkter Nachbarschaft als relevant angesehen wurden, sondern auch nicht offensichtliche Reaktionen, die nicht direkt mit der Kraftstoffsynthese verbunden sind und in anderen Bereichen des Stoffwechsels agieren, in Betracht gezogen werden müssen.

Durch eine Modellierung des gesamten Syntheseprozesses innerhalb eines genomskaligen metabolischen Netzwerkes rücken auch Reaktionen in den Fokus, die nicht unmittelbar an der Synthese des Biokraftstoffes beteiligt sind und deren Aktivität auf den ersten Blick nicht direkt ersichtlich ist. Zum Beispiel solche Reaktionen, die obligat auftretende Nebenprodukte des Syntheseweges verwenden, um sie in den Zentralstoffwechsel des Organismus zurückzuführen, wo sie noch von Nutzen sind und zum Beispiel das Zellwachstum fördern können.

Natürlich dient ein gut kuratiertes metabolisches Netzwerk nicht nur als Ausgangspunkt zu einer Modellierung mittels der Flussbilanzanalyse. Es hilft besonders bei der Sichtung und Auswahl eines optimalen Produktionsstammes für die Synthese des gewünschten Biokraftstoffes, da für die Rekonstruktion nur die annotierte Genomsequenz als Ausgangspunkt dient und diese bereits für viele phototrophe Organismen vorhanden ist.

Resümee und Ausblick

Resümee und Ausblick

Bis zu einem detaillierten Verständnis des Stoffwechsels von Cyanobakterien steht aufgrund der Komplexität dieser Thematik noch ein sehr langer Weg bevor. Eine sehr wichtige Fragestellung besteht dabei in der Klärung des grundlegenden phototrophen Wachstums dieser Organismen, aber auch die Verknüpfung des Übergangs von einem phototrophen Wachstum am Tage hin zu einem Stoffwechsel bei Abwesenheit von Licht. Dabei ist vor allem interessant, welche zeitlichen Prozesse im Stoffwechsel ablaufen und wie diese koordiniert werden.

In dieser Arbeit konnte gezeigt werden, dass eine genomskalige Modellierung des Stoffwechsels dazu beitragen kann, das grundlegende Verständnis zu fördern und wichtige Erkenntnisse in Bezug auf den cyanobakteriellen Stoffwechsel zu gewinnen. Für die Analyse eines genomskaligen metabolischen Netzwerkes hat sich hierbei die Flussbilanzanalyse als sehr hilfreiche Methode herausgestellt.

Sie ermöglicht die Vorhersage von konkreten Wachstumsraten sowie die Untersuchung der Einflüsse diverser Netzwerktopologien auf die optimale Verteilung der internen Stoffwechselflüsse. Zudem können auch komplexere Stoffwechselszenarien mit unterschiedlichen Eingangsparametern analysiert werden, wie unter anderem bei einer Simulation des Überganges vom Stoffwechsel bei Abwesenheit von Licht hin zum phototrophen Wachstum oder die Modellierung eines kompletten Tagesverlaufs.

Zudem kann ein genomskaliges Modell des Stoffwechsels eines Organismus als ein Sammelwerk für jegliches Wissen über den Stoffwechsel dienen und somit zum allgemeinen Verständnis des Stoffwechsels beitragen. Dabei können die Erkenntnisse aus der Rekonstruktion und die Ergebnisse der Modellierung aus der Flussbilanzanalyse dazu beitragen, biologische Fragen in einem systematischen Kontext zu sehen und die Einflüsse einzelner Änderungen, zum Beispiel durch eine genetische Manipulation, auf den gesamten Stoffwechsel zu betrachten. Unstimmigkeiten und Widersprüche, die sich innerhalb des Rekonstruktions- und Analyseprozesses ergeben, können tiefgreifende Diskussionen anregen und in Kombination mit den Informationen aus dem metabolischen Modell als Wissensbasis und gleichsam als Ausgangspunkt für weitere Forschungsansätze dienen, um durch weitere Analysen oder experimentelle Ansätze das Wissen über den Stoffwechsel der Cyanobakterien zu vermehren und zu verfeinern. Eine der Stärken der genomskaligen Rekonstruktion und der Flussbilanzanalyse liegt darin, dass im Vorfeld sehr wenige bis gar keine experimentellen Daten benötigt werden und nur die annotierte Genomsequenz als Ausgangsbasis dient.

Ein weiterer interessanter Punkt zu einem verbesserten Verständnis des Stoffwechsels von Cyanobakterien liegt in der Betrachtung der Diversität dieses Phylums. Denn trotz des engen Verwandtschaftsgrades innerhalb der Cyanobakterien, herrscht selbst in den zentralen Teilen des Stoffwechsels ein hohes Maß an Vielfalt und metabolischem Potenzial.

Unter diesem Aspekt könnte eine Modellierung von mehreren Cyanobakterienstämmen, in Kombination mit Daten über die Zusammensetzung der Biomasse und des Zellwachstums der unterschiedlichen Stämme, weitere Aufschlüsse über den Stoff-

wechsel der Cyanobakterien liefern. Natürlich weisen die unterschiedlichen Cyanobakterienstämme unterschiedliche Lebensrhythmen, ökologische Lebensräume und Zellformen auf. In dieser Hinsicht ist zu nennen, dass zum Beispiel diazotrophe, einzellige Cyanobakterien, wie die Gattung *Cyanothece*, andere Kriterien für ihren Tagesrhythmus beziehungsweise die zeitliche Koordinierung des Stoffwechsels aufweisen als filamentöse Cyanobakterien, die verschiedene Zelltypen ausbilden, wie unter anderem die Heterocysten bei der Gattung *Anabaena*. Aus diesen Gründen scheinen auf den ersten Blick die unterschiedlichen Stämme schwer miteinander vergleichbar zu sein, doch es ist zu bedenken, dass die meisten Cyanobakterien eine annähernd gleiche Lebensweise aufweisen. Es wäre somit eine annehmbare Herausforderung, die unterschiedlichsten Stämme auf Ebene der Stoffwechselmodellierung gegenüberzustellen und die metabolischen Abläufe innerhalb der Zelle zu vergleichen.

Eine solche umfassende Modellierung von verschiedenen Stämmen gewährt zudem einen wichtigen Ausgangspunkt für die Stammwahl bei der cyanobakteriellen Produktsynthese. In diesen Zusammenhang unterstützt eine Analyse des metabolischen Potenzials eines Stammes, zum Beispiel wie in dieser Arbeit anhand der Flussbilanzanalyse zur Untersuchung der Biokraftstoffsynthese gezeigt, dabei, die geeignete Organismus-Produkt-Kombination im Vorfeld auszuwählen und die Synthese des gewünschten Produktes zu optimieren, um zusammen eine erhöhte Produktausbeute zu erzielen. Ein solcher Schritt bietet einen großen Vorteil für die industrielle Produktion dieser Stoffe mithilfe von Cyanobakterien.

Netzwerk von *Synechocystis* sp. PCC 6803

Tabelle A.1.: Stoffwechselreaktionen für das metabolische Netzwerk von *Synechocystis* sp. PCC 6803, die zur Zeit der Rekonstruktion in Datenbanken oder der Literatur nicht verfügbar waren. Für die jeweiligen Reaktionen konnten keine spezifischen Gene im Genom von *Synechocystis* sp. PCC 6803 identifiziert werden. Die Integration dieser Reaktionen in das metabolische Netzwerk erfolgte, um die entsprechenden Stoffwechselwege zu vervollständigen.

EC-Nummer	Reaktion	Stoffwechselweg	Bemerkung
-	ATP + L-Aspartat + L-Glutamin + H ₂ O => AMP + Diphosphat + L-Asparagin + L-Glutamat	Asparagin	Es konnte kein Enzym identifiziert werden
3.1.3.27	Phosphatidylglycerophosphat + H ₂ O => Phosphatidylglycerol + Orthophosphat	Glycerolipid	Es konnte kein Enzym identifiziert werden
-	3-D-Glucosyl-1,2-Diacylglycerol <=> 1,2-Diacyl-3-β-D-Galactosyl-sn-Glycerol	Glycerolipid	unbekannte Epimerase [231]
-	3-D-Glucosyl-1,2-Diacylglycerol <=> 1,2-Diacyl-3-β-D-Galactosyl-sn-Glycerol	Glycerolipid	unbekannte Epimerase [231]
2.3.1.31	Acetyl-CoA + L-Homoserin => CoA + O-Acetyl-L-Homoserin	Methionin	Es konnte kein Stoffwechselweg für die Methioninsynthese identifiziert werden, für das Modell wurde der Stoffwechselweg von <i>Microcystis aeruginosa</i> NIES-843 übernommen
2.5.1.49	O-Acetyl-L-Homoserin + Hydrogen sulfid => L-Homocystein + Acetat	Methionin	Es konnte kein Stoffwechselweg für die Methioninsynthese identifiziert werden, für das Modell wurde der Stoffwechselweg von <i>Microcystis aeruginosa</i> NIES-843 übernommen
4.1.1.50	S-Adenosyl-L-Methionin + H ⁺ => S-Adenosylmethioninamin + CO ₂	Methionin	Es konnte kein Enzym identifiziert werden [232]
2.7.7.9	UTP + D-Glucose-1-Phosphat <=> Diphosphat + UDP-Glucose	Glucose	Es konnte keine UDP-Glucose-Phosphorylase (2.7.7.9.) identifiziert werden
3.6.1.54	UDP-2,3-Bis(3-Hydroxytetradecanoyl)glucosamin + H ₂ O => UMP + Lipid X	Lipid A-Disaccharid	Es konnte kein Enzym identifiziert werden
-	5'-Deoxyadenosin + H ₂ O => 5-Deoxy-D-Ribose + Adenin	Purin	unbekannte Nukleosidase; wird im Modell für den Abbau von 5'-Deoxyadenosin benötigt
2.5.1.16	S-Adenosylmethioninamin + Putrescin => 5'-Methylthioadenosin + Spermidin + H ⁺	Spermidin	Es konnte kein Enzym identifiziert werden [232]
2.5.1.16	S-Adenosylmethioninamin + Spermidin => 5'-Methylthioadenosin + Spermin + H ⁺	Spermidin	Es konnte kein Enzym identifiziert werden [232]
-	ATP + 4-Amino-2-Methyl-5-Phosphomethylpyrimidin => ADP + 2-Methyl-4-Amino-5-Hydroxymethylpyrimidin diphosphat	Thiamin-Pyrophosphat	Es konnte kein Enzym identifiziert werden
1.2.1.72	D-Erythrose-4-Phosphat + NAD ⁺ + H ₂ O => 4-Phospho-D-Erythronat + NADH + H ⁺	Vitamin B ₆	Es konnte kein Enzym identifiziert werden [233]
1.1.1.290	4-Phospho-D-Erythronat + NAD ⁺ => 2-Oxo-3-Hydroxy-4-Phosphobutanoat + NADH + H ⁺	Vitamin B ₆	Es konnte kein Enzym identifiziert werden [233]
2.6.1.52	O-Phospho-4-Hydroxy-L-Threonin + 2-Oxoglutarat <=> 2-Oxo-3-Hydroxy-4-Phosphobutanoat + L-Glutamat	Vitamin B ₆	Es konnte kein Enzym identifiziert werden [233]
1.14.13.83	Precorrin 3A + Oxygen + NADH + H ⁺ => Precorrin 3B + NAD ⁺ + H ₂ O	Vitamin B ₁₂	Es konnte kein Enzym für eine Precorrin-3B-Synthase identifiziert werden [234]

Tabelle A.2.: Stoffwechselreaktionen für das metabolische Netzwerk von *Synechocystis* sp. PCC 6803, die zur Zeit der Rekonstruktion in Datenbanken nicht verfügbar waren. Die Identifizierung dieser Reaktionen erfolgte über die entsprechende Literatur oder über eine BLAST-Suche. Die Beleg-Spalte gibt jeweils die Art der Informationsquelle an: B) Beleg direkt in der Primärliteratur, C) das entsprechende Gen konnte mittels einer BLAST-Suche identifiziert werden, D) ungeklärtes Enzym mit potenzieller metabolischer Funktion oder Funktion von einem anderen Organismus übertragen.

Gen	EC-Nummer	Reaktion	Stoffwechselweg	Beleg	Bemerkung
sll1336	3.5.3.6	L-Arginin + H ₂ O => L-Citrullin + NH ₃	Arginin	B	[235]
slr1022	2.6.1.13	L-Ornithin + 2-Oxoglutarat <=> L-Glutamat 5-Semialdehyd + L-Glutamat	Arginin	B	[235]
slr0782	1.4.3.4	L-Arginin + PQ + H ₂ O => 5-Guanidino-2-Oxopentanoat + NH ₃ + PQH ₂	Arginin	B	Arginin-Dehydrogenase [236]
slr0782	1.4.3.4	L-Arginin + PQ + H ₂ O => 5-Guanidino-2-Oxopentanoat + NH ₃ + PQH ₂	Arginin	B	Arginin-Dehydrogenase [236]
-	-	5-Guanidino-2-Oxopentanoat + H ₂ O ₂ => 4-Guanidinobutanoat + CO ₂ + H ₂ O	Arginin	B	spontan [235]
sll1077 oder sll0228	3.5.3.7	4-Guanidinobutanoat + H ₂ O => 4-Aminobutanoat + Urea	Arginin	B	[235]
sll0428 oder sll0191	1.14.11.2	L-Prolin + 2-Oxoglutarat + Oxygen => trans-4-Hydroxy-L-Prolin + Succinat + CO ₂	Prolin	C	BLAST: sll0428 57 % Ident., 257 Score, E-Wert 8 e ⁻⁷⁰ und sll0191 48 % Ident., 197 Score, E-Wert 8 e ⁻⁵² Prolyl-Hydroxylase von <i>Trichodesmium erythraeum</i> IMS101
sll1713	4.1.1.81	L-Threonin O-3-Phosphat => D-1-Aminopropan-2-ol O-Phosphat + CO ₂	Threonin	C	BLAST: Ident., 57.5 %, Score, 414, E-Wert e ⁻¹¹⁷ zu Threonin-Phosphat-Decarboxylase von <i>Anabaena variabilis</i> ATCC 29413
sll1237	1.3.3.4	2 Protoporphyrinogen IX + 3 Oxygen <=> 2 Protoporphyrin + 6 H ₂ O	Chlorophyll	C	BLAST: Ident., 54%, Score, 308, E-Wert 4e ⁻⁸⁵ zu Protoporphyrinogen-Oxidase von <i>Anabaena</i> sp. PCC 7120
slr1923	1.3.1.75	Divinylprotochlorophyllid + NADPH + H ⁺ => Protochlorophyllid + NADP ⁺	Chlorophyll	B	[237], [238]
slr1979	2.6.1.85	Chorismat + L-Glutamin => 4-Amino-4-Deoxychorismat + L-Glutamat	Folat	C	BLAST: 63.5 % Ident., 574 Score, E-Wert e ⁻¹⁶⁵ zu Para-Aminobenzoat-Synthase-Komponente I von <i>Cyanosyce</i> sp. PCC 8801
slr2141	1.5.1.20	5-Methyltetrahydrofolat + NADP ⁺ <= 5,10-Methylenetetrahydrofolat + NADPH + H ⁺	Folat	C	BLAST: 75 % Ident., 450 Score, E-Wert e ⁻¹²⁸ zu Methylenetetrahydrofolat-Reduktase von <i>Cyanosyce</i> sp. PCC 8801
slr0091	1.2.1.3	Glycolaldehyd + NAD ⁺ + H ₂ O => Glycolat + NADH + H ⁺	Folat	D	könnte durch eine Aldehyd-Dehydrogenase 1.2.1.21 katalysiert werden
slr2026	1.5.1.3	Tetrahydrofolat + NADP ⁺ <= Dihydrofolat + NADPH + H ⁺	Folat	D	slr2026 kodiert für FoIP (2.5.1.15), welche als eine Dihydropteroat-Reduktase (1.5.1.3) fungieren kann (in <i>H. Pylori</i>) [239]
slr2141	1.5.1.20	5-Methyltetrahydrofolat + NAD ⁺ <= 5,10-Methylenetetrahydrofolat + NADH + H ⁺	Folat	C	BLAST: 75 % Ident., 450 Score, E-Wert e ⁻¹²⁸ zu Methylenetetrahydrofolat-Reduktase von <i>Cyanosyce</i> sp. PCC 8801
slr0600, -slr1562, -ssr2061	-	2 Glutathion + NADP ⁺ <=> Glutathion disulfid + NADPH + H ⁺	Allgemein	B	[240]
sll1973, -slr1510, -sll1848	-	1.2759 sn-Glycerol 3-Phosphat + 1.3302 Hexadecanoyl-[acp] + 0.0774 Hexadecanoyl-[acp] + 0.0034 Octadecanoyl-[acp] + 0.0635 Oleoyl-[acyl-carrier-Protein] + 0.4236 Linoleoyl-[acp] + 0.6315 gamma-Linolenoyl-[acp] + 0.0032 alpha-Linolenoyl-[acp] + 0.0190 Stearidonoyl-[acp] => Phosphatidat + 2.5518 Acyl-carrier-Protein	Glycerolipid	-	kombinierte Reaktion von 2.3.1.15 und 2.3.1.51
sll1973, -slr1510, -sll1848	-	1.2954 sn-Glycerol 3-Phosphat + 1.3061 Hexadecanoyl-[acp] + 0.0965 Hexadecanoyl-[acp] + 0.0223 Octadecanoyl-[acp] + 0.0584 Oleoyl-[acyl-carrier-Protein] + 0.3334 Linoleoyl-[acp] + 0.7253 gamma-Linolenoyl-[acp] + 0.0095 alpha-Linolenoyl-[acp] + 0.0393 Stearidonoyl-[acp] => Phosphatidat + 2.5908 Acyl-carrier-Protein	Glycerolipid	-	kombinierte Reaktion von 2.3.1.15 und 2.3.1.51
sll1973, -slr1510, -sll1848	-	1.3747 sn-Glycerol 3-Phosphat + 1.5612 Hexadecanoyl-[acp] + 0.1804 Hexadecanoyl-[acp] + 0.0434 Octadecanoyl-[acp] + 0.4011 Oleoyl-[acyl-carrier-Protein] + 0.5331 Linoleoyl-[acp] + 0.0228 gamma-Linolenoyl-[acp] + 0.0165 alpha-Linolenoyl-[acp] => Phosphatidat + 2.7494 Acyl-carrier-Protein	Glycerolipid	-	kombinierte Reaktion von 2.3.1.15 und 2.3.1.51
sll1973, -slr1510, -sll1848	-	1.32535 sn-Glycerol 3-Phosphat + 1.3327 Hexadecanoyl-[acp] + 0.0227 Hexadecanoyl-[acp] + 0.0457 Octadecanoyl-[acp] + 0.1701 Oleoyl-[acyl-carrier-Protein] + 0.8637 Linoleoyl-[acp] + 0.2158 gamma-Linolenoyl-[acp] => Phosphatidat + 2.6507 Acyl-carrier-Protein	Glycerolipid	-	kombinierte Reaktion von 2.3.1.15 und 2.3.1.51
sll0545	3.1.3.4	Phosphatidat + H ₂ O => 1,2-Diacyl-sn-Glycerol + Orthophosphat	Glycerolipid	B	[241]

Anhang A. Netzwerk von *Synechocystis* sp. PCC 6803

slr0545	3.1.3.4	Phosphatidat + H ₂ O => 1,2-Diacyl- <i>sn</i> -Glycerol + Orthophosphat	Glycerolipid	B	[241]
slr0545	3.1.3.4	Phosphatidat + H ₂ O => 1,2-Diacyl- <i>sn</i> -Glycerol + Orthophosphat	Glycerolipid	B	[241]
slr0384	2.4.1.-	UDP-6-Sulfoquinovosdiacylglycerol + 1,2-Diacyl- <i>sn</i> -Glycerol => Sulfoquinovosdiacylglycerol + UDP	Glycerolipid	B	[242]
slr1377	2.4.1.157	UDP-Glucose + 1,2-Diacyl- <i>sn</i> -Glycerol => UDP + 3-D-Glucosyl-1,2-Diacylglycerol	Glycerolipid	B	[231]
slr1377	2.4.1.157	UDP-Glucose + 1,2-Diacyl- <i>sn</i> -Glycerol => UDP + 3-D-Glucosyl-1,2-Diacylglycerol	Glycerolipid	B	[231]
slr1538	2.1.1.152	Precorrin 6X + Acetat + <i>S</i> -Adenosyl-L-Homocystein <=> Precorrin 5 + <i>S</i> -Adenosyl-L-Methionin + H ₂ O	Vitamin B ₁₂	C	BLAST: Ident., 57.1 %, Score, 356, E-Wert e ⁻¹⁰⁴ zu Cobalt-Precorrin-6A-Synthase von <i>Trichodesmium erythraeum</i> IMS101 slr0895 kodiert für CysQ, welches in vielen Organismen als 3.1.3.7 fungiert
slr0895	3.1.3.7	Adenosin 3',5'-Bisphosphat + H ₂ O => AMP + Orthophosphat	Sulfat	D	
slr1908	1.1.1.81	D-Glycerat + NADP ⁺ <=> Hydroxypyruvat + NADPH + H ⁺	Glyoxylat	B	[119]
slr0229	1.1.1.60	D-Glycerat + NADP ⁺ <=> Tartronat-Semialdehyd + NADPH + H ⁺	Glyoxylat	B	[119]
slr1981	4.1.1.47	2 Glyoxylat => Tartronat-Semialdehyd + CO ₂	Glyoxylat	B	[119]
slr1556	1.2.3.5	Glyoxylat + Oxygen + H ₂ O => Oxalat + H ₂ O ₂	Glyoxylat	B	[119]
slr0006	2.6.1.4	Glycin + 2-Oxoglutarat <=> Glyoxylat + L-Glutamat	Glyoxylat	B	[119]
slr0633	1.4.3.19	Glycin + H ₂ O + Oxygen => Glyoxylat + NH ₃ + H ₂ O ₂	Glyoxylat	C	BLAST: Ident., 70.6 %, Score, 936, E-Wert 0.0 zu Glycin-Oxidase von <i>Cyanospora</i> sp. PCC 8801
slr0895	3.1.3.7	3'-Phosphoadenylyl sulfat + H ₂ O => Adenylyl sulfat + Orthophosphat	Sulfat	D	slr0895 kodiert für CysQ, welches in vielen Organismen als 3.1.3.7 fungiert
slr0990	6.3.2.2	ATP + L-Glutamat + L-Cystein => ADP + Orthophosphat + gamma-L-Glutamyl-L-Cystein	Glutathion	C	BLAST Search: 68 % Ident., 540 Score, E-Wert e ⁻¹⁵⁵ zu Glutamat-Cystein-Ligase von <i>Cyanospora</i> sp. PCC 8801
slr1124	3.1.3.3	O-Phospho-L-Serin + H ₂ O => L-Serin + Orthophosphat	Serin	B	BLAST: Ident., 31 %, Score, 102, E-Wert 2e ⁻²³ zu PspA von <i>Hydrogenobacter thermophilus</i> [131]
slr1211	6.6.1.2	Hydrogenobyrinat a,c diamid + Cobalt ion + ATP + H ₂ O => Cob(II)yrinat a,c diamid + Orthophosphat + ADP + H ⁺	Vitamin B ₁₂	D	slr2026 kodiert für FoIP (2.5.1.15), welche als eine Dihydropteroat-Reduktase (1.5.1.3) fungieren kann (in <i>H. Pylori</i>) [239]
slr0953	3.1.3.24	Sucrose-6'-Phosphat + H ₂ O => Sucrose + Orthophosphat	Zucker	B	[243]
slr0084	3.1.3.15	L-Histidinol-Phosphat + H ₂ O => L-Histidinol + Orthophosphat	Histidin	C	BLAST: 56 % Ident., 189 Score, E-Wert e ⁻⁴⁹ zu <i>hisB</i> von <i>Cyanospora</i> sp. ATCC 51142
slr1564 oder slr0186	2.3.1.182	Acetyl-CoA + Pyruvat + H ₂ O => (R)-2-Methylmalat + CoA	Pyruvat	B	[224]
slr1722	5.5.1.4	D-Glucose-6-Phosphat => 1D-Myo-Inositol 3-Phosphat	Inositol	B	[244]
slr1981	5.5.1.4	D-Glucose-6-Phosphat => 1D-Myo-Inositol 3-Phosphat	Hem	B	[245]
slr1959	3.1.3.25	1D-Myo-Inositol 3-Phosphat + H ₂ O => myo-Inositol + Orthophosphat	Inositol	B	[246]
slr0541	1.14.19.1	Hexadecanoyl-[acp] + NADPH + H ⁺ + Oxygen <=> Hexadecenoyl-[acp] + NADP ⁺ + 2 H ₂ O	Lipid	B	<i>desC</i> , Δ ⁹ -Desaturase [247, 248]
slr0541	1.14.19.1	Octadecanoyl-[acp] + NADPH + H ⁺ + Oxygen <=> Oleoyl-[acp] + NADP ⁺ + 2 H ₂ O	Lipid	B	<i>desC</i> , Δ ⁹ -Desaturase [247, 248]
slr1508	2.4.1.184	2 1,2-Diacyl-3-β-D-Galactosyl- <i>sn</i> -Glycerol => Digalactosyl-Diacylglycerol + 1,2-Diacyl- <i>sn</i> -Glycerol	Glycerolipid	B	[249]
slr0262	1.14.19.3	Linoleoyl-[acp] + NADPH + H ⁺ + Oxygen <=> gamma-Linolenoyl-[acp] + NADP ⁺ + 2 H ₂ O	Lipid	B	<i>desD</i> , Δ ⁶ -Desaturase [247, 248]
slr1441	1.14.19.-	Linoleoyl-[acp] + NADPH + H ⁺ + Oxygen <=> alpha-Linolenoyl-[acp] + NADP ⁺ + 2 H ₂ O	Lipid	B	<i>desB</i> , Δ ¹⁵ -Desaturase [247, 248]
slr0262	1.14.19.3	alpha-Linolenoyl-[acp] + NADPH + H ⁺ + Oxygen <=> Stearidonoyl-[acp] + NADP ⁺ + 2 H ₂ O	Lipid	B	<i>desD</i> , Δ ⁶ -Desaturase [247, 248]
slr1363	1.1.1.86	(R)-Pantoat + NADP ⁺ <=> 2-Dehydropantoat + NADPH + H ⁺	Pantothenat	D	slr1363 kodiert für IlvC (1.1.1.86), welches auch als 2-Dehydropantoat-2-Reduktase fungieren kann (in <i>C. glutamicum</i>) [250]
slr1833	2.4.1.129	Undecaprenyl-Diphospho-N-Acetylmuramoyl-(N-Acetylglucosamine)-L-Alanyl-D-Glutamyl-Meso-2,6-Diaminopimeloyl-D-Alanyl-D-Alanin + H ₂ O => di-Trans,poly-cis-Undecaprenyl diphosphat + [GlcNAc-(1->4)-Mur2Ac(oil)-L-Ala-G-D-Glu-A2pm-D-Ala-D-Ala)	Peptidoglycan	C	BLAST: Ident., 56%, Score, 670, E-Wert 0.0 zu Peptidoglycan-Glycosyltransferase von <i>Cyanospora</i> sp. PCC 7424
slr0529	4.1.2.22	D-Fructose-6-Phosphat + Orthophosphat => Acetyl-Phosphat + D-Erythrose-4-Phosphat + H ₂ O	OPP	C	BLAST: Ident., 80.4 %, Score, 1203, E-Wert 0.0 zu D-Xylulose-5-Phosphat/D-Fructose-6-Phosphat-Phosphoketolase von <i>Cyanospora</i> sp. PCC 8801
slr0453	4.1.2.22	D-Fructose-6-Phosphat + Orthophosphat => Acetyl-Phosphat + D-Erythrose-4-Phosphat + H ₂ O	OPP	C	BLAST: Ident., 67 %, Score, 1163, E-Wert 0.0 zu D-Xylulose-5-Phosphat/D-Fructose-6-Phosphat-Phosphoketolase von <i>Acaryochloris marina</i> MBIC11017
slr0529	4.1.2.22	D-Xylulose-5-Phosphat + Orthophosphat => Acetyl-Phosphat + D-Glyceraldehyd 3-Phosphat + H ₂ O	OPP	C	BLAST: Ident., 80.4 %, Score, 1203, E-Wert 0.0 zu D-Xylulose-5-Phosphat/D-Fructose-6-Phosphat-Phosphoketolase von <i>Cyanospora</i> sp. PCC 8801
slr0453	4.1.2.22	D-Xylulose-5-Phosphat + Orthophosphat => Acetyl-Phosphat + D-Glyceraldehyd 3-Phosphat + H ₂ O	OPP	C	BLAST: Ident., 67 %, Score, 1163, E-Wert 0.0 zu D-Xylulose-5-Phosphat/D-Fructose-6-Phosphat-Phosphoketolase von <i>Acaryochloris marina</i> MBIC11017

slr1350	1.14.19.-	Oleoyl-[acp] + NADPH + H ⁺ + Oxygen <=> Linoleoyl-[acp] + NADP ⁺ + 2 H ₂ O	Lipid	B	desA, Δ ¹² -Desaturase [247, 248]
-	-	1 Photon + 1 P700 + 1 Oxygen => 1 P700p + 1 Superoxid anion	Algemein	-	Mehler-Reaktion am PSI
slr0550, slr1521	-	2 NADPH + 2 H ⁺ + Oxygen => 2 NADP ⁺ + 2 H ₂ O	Algemein	-	Mehler-ähnliche-Reaktion [150]
-	-	1 Oxygen + 1 QAn => 1 Superoxid anion + 1 QA	Algemein	-	ROS-Entwicklung am PSII
slr0741	1.2.7.1	2 Reduced ferredoxin + Acetyl-CoA + CO ₂ + 2 H ⁺ <=> 2 Oxidized ferredoxin + Pyruvat + CoA	Pyruvat und TCA	C	BLAST: 77 % Ident., 1940 Score, E-Wert 0,00 zu Pyruvat-Flavodoxin-Oxidoreduktase von <i>Micorystis aeruginosa</i> [119]
slr1359	1.2.2.1	Methansäure + 2 Ferricytochrom b1 => CO ₂ + 2 Ferrocycytochrom b1 + 2 H ⁺	Glyoxylat	B	
slr1981	4.1.1.71	2-Oxoglutarat => Succinat semialdehyd + CO ₂	Pyruvat und TCA	C	BLAST Search: Ident., 80%, Score, 911, E-Wert 0.0 zu 2-Oxoglutarat-Decarboxylase von <i>Synechococcus</i> sp. PCC 7002 (SYNPCC7002_A2770) [142]
slr1022	2.6.1.19	4-Aminobutanoat + 2-Oxoglutarat <=> Succinat semialdehyd + L-Glutamat	Pyruvat und TCA	B	[235]
slr0861	2.1.2.2	5'-Phosphoribosylglycinamid + Formate + ATP => 5'-Phosphoribosyl-N-Formylglycinamid + ADP + Orthophosphat	Purin	B	PurT-Reaktion von GAR-Transformylase [251]
slr1797	4.1.3.40	4-Hydroxybenzoat + Pyruvat <=> Chorismat	Chinon und Tocopherol	B	BLAST: Ident., 46 %, Score, 162, E-Wert 6e ⁻⁴² zu 4-Hydroxybenzoat-Synthetase von <i>Prochlorococcus marinus</i> str. MIT 9303 [129]
slr0926	2.5.1.39	all-Trans-Nonaprenyl diphosphat + 4-Hydroxybenzoat => Diphosphat + 3-Nonaprenyl-4-Hydroxybenzoat	Chinon und Tocopherol	B	PlqA [129]
slr0936 oder slr1099 oder slr1300	-	3-Nonaprenyl-4-Hydroxybenzoat + 4 H ⁺ => 2-Methyl-6-Solanyl-1,4-benzoquinol + H ₂ O	Chinon und Tocopherol	D	Unbekannte Reaktion, kombinierte Schritte aus der Plastochinonsynthese [129]
slr0400	3.1.3.-	5-Amino-6-(5'-Phosphoribitylamino)uracil + H ₂ O <=> 5-Amino-6-(1-D-ribitylamino)uracil + Orthophosphat	Riboflavin	D	unklar, slr0400 Kandidaten-Gen
slr0945	1.5.1.38	Reduced FMN + NADP ⁺ <=> FMN + NADPH + H ⁺	Riboflavin	C	BLAST: Ident., 87.6 %, Score, 363, E-Wert e ⁻¹⁰² zu NADPH-abhängige-FMN-Reduktase von <i>Cyanosche</i> sp. ATCC 51142 [252]
slr1790	1.3.3.4	2 Protoporphyrinogen IX + 3 Oxygen <=> 2 Protoporphyrin + 6 H ₂ O	Chlorophyll	B	
slr0118	-	Aminoimidazol-Ribotid + S-Adenosyl-L-Methionin <=> 4-Amino-2-Methyl-5-Phosphomethylpyrimidin + L-Methionin + 5'-Deoxyadenosin + Methansäure + Kohlenmonoxid	Thiamin-Pyrophosphat	D	ThiC, ThiD nicht annotiert, kombinierte Reaktion
slr0102, slr0633	-	L-Cystein + ATP + 1-Deoxy-D-Xylulose-5-Phosphat + L-Tyrosin + S-Adenosyl-L-Methionin + NADPH + 3 H ⁺ => 4-Methyl-5-(2-Phosphoethyl)-Thiazol + Diphosphat + L-Methionin + L-Alanin + 4-	Thiamin-Pyrophosphat	D	BLAST: slr0102 Ident., 68 %, Score, 98, E-Wert 2e ⁻²² zu ThiS von <i>Nostoc punctiforme</i> ATCC 29133, kombinierte Reaktion mit ThiG
slr0401	-	ATP + H ₂ O + Putrescin => ADP + Orthophosphat + Putrescin	Transport	B	[253, 254]
slr0401	-	ATP + H ₂ O + Spermidin => ADP + Orthophosphat + Spermidin	Transport	B	[253, 254]
slr1719 oder slr0001	1.14.99.40	Reduced FMN + Oxygen + NADH + H ⁺ => Dimethylbenzimidazol + D-Erythrose-4-Phosphat + H ₂ O + NAD ⁺	Riboflavin	D	BLAST: slr1719 Ident., 67 %, Score, 278, E-Wert 3e ⁻⁷⁶ zu Nitro-Reduktase von <i>Trichodesmium erythraeum</i> IMS101
-	-	NADPH + 5 H ⁺ + PQ => NADP ⁺ + PQH ₂ + 4 H ⁺	Algemein	-	Zyklischer Elektronenfluss am PS [255-257]
slr0001	1.16.8.1	2 Cob(II)yrinat a,c diamid + Reduced FMN <=> 2 Cob(II)yrinat a,c diamid + FMN + 2 H ⁺	Vitamin B ₁₂	C	BLAST: Ident., 31 %, Score, 57, E-Wert 4e ⁻¹⁰ zu 4-Hydroxyphenylacetat-3-Monooxygenase von <i>B. melitensis</i> , welches als Corrin-Reduktase fungieren kann [258]
slr1500	2.4.2.21	Nicotinat D-ribonucleotid + Dimethylbenzimidazol => Nicotinat + N1-(5-Phospho-Alpha-D-ribosyl)-5,6-Dimethylbenzimidazol + H ⁺	Vitamin B ₁₂	C	BLAST: Ident., 61 %, Score, 439, E-Wert e ⁻¹²⁴ zu Nicotinat-Nucleotid-Dimethylbenzimidazol-Phosphoribosyltransferase von <i>Cyanosche</i> sp. PCC 8801
slr1124	3.1.3.73	N1-(5-Phospho-Alpha-D-ribosyl)-5,6-Dimethylbenzimidazol + H ₂ O => alpha-Ribazol + Orthophosphat	Vitamin B ₁₂	C	BLAST: Ident., 46 %, Score, 397, E-Wert e ⁻¹¹² zu Ribazol-5'-Phosphatase von <i>Synechococcus</i> sp. CC9902

Tabelle A.3.: Erläuterung der zusätzlichen Abkürzungen, die in den Stoffwechselabbildungen in dieser Arbeit verwendet werden.

zusätzliche Abkürzungen			
2OG	2-Oxoglutarat	Leu	Leucin
2Oiv	2-Oxoisovalerat	LpAD	Lipid A-Disaccharid
2PGL	2-Phosphoglycolat	Lys	Lysin
3HbCoA	Hydroxy-Butyryl-CoA	MaCoA	Malonyl-CoA
3PG	3-Phospho-Glycerol-Phosphat	Mal	Malat
3P-HP	3-Phospho-Hydroxypyruvat	mDom	<i>meso</i> -2,6-Diaminopimelat
4AB	4-Aminobutanoat	Met	Methionin
5PrPP	5-Phospho-Ribose-1-Diphosphat	MGDG	Monogalactosyl-Diacyl-Glycerol
6PG	6-Phospho-Gluconat	MTH	Methyltetrahydrofolat
A4Sa	Aspartat-4-Semialdehyd	NAD	Nicotinamid-Adenin-Dinukleotid
AcAh	Acetaldehyd	NADP	Nicotinamid-Adenin-Dinukleotid-Phosphat
AcCoA	Acetyl-CoA	NDH	NADPH-Dehydrogenase
Ace	Acetat	OA	Oxalacetat
AceP	Acetyl-Phosphat	P5C	Pyrrolin-5-Carboxylat
ACP	Acyl-Carrier-Protein	PCP	Plastocyanin
ADP	Adenosin-5'-Diphosphat	PEP	Phosphoenolpyruvat
ADPG	ADP-Glucose	PepGlc	Peptidoglycan
AIR	Aminoimidazol-Ribotid	PG	Phosphatidyl-Glycerol
Ala	Alanin	PG2	Glycerat-2-Phosphat
Arg	Arginin	PG3	Glycerat-3-Phosphat
Asn	Asparagin	Phe	Phenylalanin
Asp	Aspartat	Pho	Phosphat
ATP	Adenosin-5'-Triphosphat	Phyq	Phyllochinon
ATPase	ATP-Synthase	Ppg	Protoporphyrinogen
BCoA	Buyryl-CoA	PPP	Phytlyl-Diphosphat
ButAc	Butansäure	PQ	Plastoquinon
ButAh	Butyraldehyd	Prep	Prephenat
bCaro	β -Carotin	Pro	Prolin
Chlp	Chlorophyll <i>a</i>	PS I	Photosystem I
Chor	Chorismat	PS II	Photosystem II
Cit	Citrat	Ptd	Phosphatidat
COX	Cytochrom- <i>c</i> -Oxidase	Ptsn	Putrescin
CrCoA	Crotonoyl-CoA	Pyr	Pyruvat
Cys	Cystein	R5P	Ribose-5-Phosphat
Cyt b6f	Cytochrom- <i>b</i> _{6f} -Komplex	Rbfv	Riboflavin
DGDG	Digalactosyl-Diacyl-Glycerol	Ru5P	Ribulose-5-Phosphat
DHAP	Dihydroxyaceton-Phosphat	RuBP	Ribulose-1,5-Bisphosphat
DmPP	Dimethylallyl-Diphosphat	S7P	Sedoheptulose-7-Phosphat
DX5P	1-Deoxy-Xylulose-5-Phosphat	SAM	S-Adenosylmethioninamin
E4P	Erythrose-4-Phosphat	SBP	Sedoheptulose-1,7-Bisphosphat
Echi	Echinenon	ScCoA	Succinyl-CoA
F6P	Fructose-6-Phosphat	SDH	Succinat-Dehydrogenase
FAD	Flavin-Adenin-Dinukleotid	Ser	Serin
FBP	Fructose-1,6-Bisphosphat	Spmid	Spermidin
FNR	Ferredoxin-NADP-Reduktase	SPP	Solanyl-Diphosphat
Fum	Fumarat	SQDG	Sulfoquinovosyldiacyl-Glycerol
G1P	Glucose-1-Phosphat	Ssa	Succinat-Semialdehyd
G6P	Glucose-6-Phosphat	Suc	Succinat
GAP	Glyceraldehyd-3-Phosphat	Sul	Sulfur

gCaro	γ -Carotin	THF	Tetrahydrofolat
GgPP	Geranylgeranyl-Diphosphat	ThPP	Thiamin-Diphosphat
GL	Glycolat	Thr	Threonin
GL6P	6-Phospho-Gluconolacton	Toco	Tocopherol
Gln	Glutamin	Trp	Tryptophan
Glu	Glutamat	TSDH	Tartronat-Semialdehyd
GLX	Glyxoylat	Tyr	Tyrosin
Gly	Glycin	Upg III	Uroporphyrinogen III
GSH	Glutathion	UpPP	Undecaprenyl-Diphosphat
His	Histidin	Val	Valin
HP	Hydroxypyruvat	Vit B6	Vitamin B ₆
Hser	Homoserin	Vit B12	Vitamin B ₁₂
Icit	Isocitrat	X5P	Xylulose-5-Phosphat
Ile	Isoleucin	Zea	Zeaxanthin
IpPP	Isopentenyl-Diphosphat		

Tabelle A.4.: Beschreibung der Reaktionen, die eine abfallenden Flussaktivität bei der Variabilitätsanalyse für die Synthese von Ethylen aufweisen (Abbildung 5.4).

Reaktions- nummer	Gen	EC- Nummer	Reaktion	Stoffwechselweg
R464	sll1182, sll1316, sll1317, slr0342, slr0343, slr1185, sml0004, smr0003, smr0010	1.10.99.1	PQH2 + 1 Oxidized rieske_B6 => PQH_B6_L + 1 Reduced rieske_B6 + 1 H ⁺	Cyt b6f (Q-Zyklus)
R465	sll1182, sll1316, sll1317, slr0342, slr0343, slr1185, sml0004, smr0003, smr0010	1.10.99.1	1 Reduced rieske_B6 + 1 Oxidized plastocyanin => 1 Oxidized rieske_B6 + 1 Reduced plastocyanin	Cyt b6f (Q-Zyklus)
R466	sll1182, sll1316, sll1317, slr0342, slr0343, slr1185, sml0004, smr0003, smr0010	1.10.99.1	PQH_B6_L + 1 HB3p_B6 => PQ_B6_S + 1 HB2p_B6 + 1 H ⁺	Cyt b6f (Q-Zyklus)
R467	sll1182, sll1316, sll1317, slr0342, slr0343, slr1185, sml0004, smr0003, smr0010	1.10.99.1	PQ_B6_S => PQ	Cyt b6f (Q-Zyklus)
R468	sll1182, sll1316, sll1317, slr0342, slr0343, slr1185, sml0004, smr0003, smr0010	1.10.99.1	PQ_B6_S + 1 HB2p_B6 + 1 H ⁺ => PQH_B6_S + 1 HB3p_B6	Cyt b6f (Q-Zyklus)
R469	sll1182, sll1316, sll1317, slr0342, slr0343, slr1185, sml0004, smr0003, smr0010	1.10.99.1	PQH_B6_S + 1 HB2p_B6 + 1 H ⁺ => PQH2 + 1 HB3p_B6	Cyt b6f (Q-Zyklus)
R492	sll1327, sll1323, sll1322, ssl2615, sll1324, sll1325, sll1326, slr1329, slr1330	3.6.3.14	14 H ⁺ + 3 ADP + 3 Orthophosphat => 3 ATP + 3 H ₂ O + 14 H ⁺	ATPase
R670	-	-	1 O ₂ _ext <=> 1 O ₂ _cyt	Transport

%recalctypearea

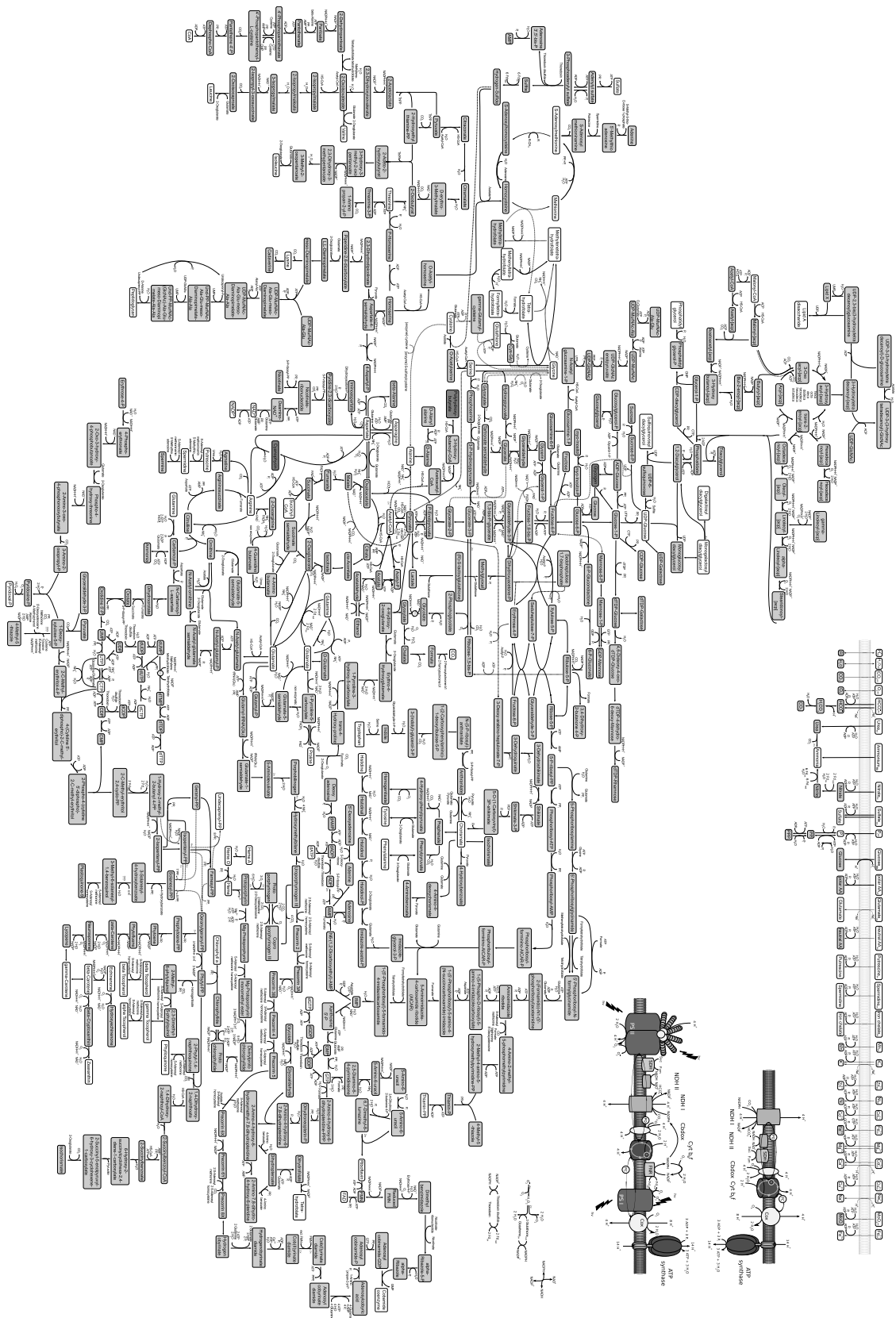


Abbildung A.1.: Stoffwechselkarte des Netzwerkes von *Synechocystis* sp. PCC 6803 (Abbildung aus Knoop et al. (2013) [94])

Literaturverzeichnis

- [1] S. B. Hedges, H. Chen, S. Kumar, D. Y. Wang, A. S. Thompson et al. (2001). A genomic timescale for the origin of eukaryotes. *BMC Evolutionary Biology*, 1(1):4
- [2] J. W. Schopf (1993). Microfossils of the Early Archean Apex Chert: New Evidence of the Antiquity of Life. *Science*, 260(5108):640–646
- [3] T. Cavalier-Smith (1987). The Simultaneous Symbiotic Origin of Mitochondria, Chloroplasts, and Microbodies. *Annals of the New York Academy of Sciences*, 503(1):55–71
- [4] B. A. Whitton und M. Potts (2000). *The Ecology of Cyanobacteria: Their Diversity in Time and Space*. Springer
- [5] C. v. d. Hoek, H. M. Jahns und D. G. Mann (1993). *Algen*. Thieme, Stuttgart [etc.]
- [6] H. Sielaff, G. Christiansen und T. Schwecke (2006). Natural products from cyanobacteria: Exploiting a new source for drug discovery. *IDrugs: the investigational drugs journal*, 9(2):119–127
- [7] L. T. Tan (2007). Bioactive natural products from marine cyanobacteria for drug discovery. *Phytochemistry*, 68(7):954–979
- [8] S. Atsumi, W. Higashide und J. C. Liao (2009). Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat Biotech*, 27(12):1177–1180
- [9] M.-D. Deng und J. R. Coleman (1999). Ethanol Synthesis by Genetic Engineering in Cyanobacteria. *Applied and Environmental Microbiology*, 65(2):523–528
- [10] P. Lindberg, S. Park und A. Melis (2010). Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metabolic Engineering*, 12(1):70–79
- [11] A. L. Mascarelli (2009). Gold rush for algae. *Nature News*, 461(7263):460–461
- [12] R. Y. Stanier, R. Kunisawa, M. Mandel und G. Cohen-Bazire (1971). Purification and properties of unicellular blue-green algae (order *Chroococcales*). *Bacteriological Reviews*, 35(2):171–205
- [13] T. Kaneko, S. Sato, H. Kotani, A. Tanaka, E. Asamizu et al. (1996). Sequence Analysis of the Genome of the Unicellular Cyanobacterium *Synechocystis* sp. Strain PCC6803. II. Sequence Determination of the Entire Genome and Assignment of Potential Protein-coding Regions. *DNA Research*, 3(3):109–136

- [14] C. A. Kerfeld, M. R. Sawaya, S. Tanaka, C. V. Nguyen, M. Phillips et al. (2005). Protein structures forming the shell of primitive bacterial organelles. *Science (New York, N.Y.)*, 309(5736):936–938
- [15] G. Grigorieva und S. Shestakov (1982). Transformation in the cyanobacterium *Synechocystis* sp. 6803. *FEMS Microbiology Letters*, 13(4):367–370
- [16] S. L. Anderson und L. McIntosh (1991). Light-activated heterotrophic growth of the cyanobacterium *Synechocystis* sp. strain PCC 6803: a blue-light-requiring process. *J. Bacteriol.*, 173(9):2761–2767
- [17] L. B. Smart, S. L. Anderson und L. McIntosh (1991). Targeted genetic inactivation of the photosystem I reaction center in the cyanobacterium *Synechocystis* sp. PCC 6803. *The EMBO journal*, 10(11):3289–3296
- [18] W. A. Kratz und J. Myers (1955). Photosynthesis and Respiration of Three Blue-Green Algae. 12. *Plant Physiology*, 30(3):275–280
- [19] C. Nägeli (1849). *Gattungen einzelliger Algen: physiologisch und systematisch Bearbeitet*. Friedrich Schulthess
- [20] C. C. Zhang, R. Jeanjean und F. Joset (1998). Obligate phototrophy in cyanobacteria: more than a lack of sugar transport. *FEMS microbiology letters*, 161(2):285–292
- [21] Y. Chen, C. K. Holtman, R. D. Magnuson, P. A. Youderian und S. S. Golden (2008). The complete sequence and functional analysis of pANL, the large plasmid of the unicellular freshwater cyanobacterium *Synechococcus elongatus* PCC 7942. *Plasmid*, 59(3):176–192
- [22] C. Sugita, K. Ogata, M. Shikata, H. Jikuya, J. Takano et al. (2007). Complete nucleotide sequence of the freshwater unicellular cyanobacterium *Synechococcus elongatus* PCC 6301 chromosome: gene content and organization. *Photosynthesis research*, 93(1-3):55–67
- [23] B. D. Rae, B. M. Long, M. R. Badger und G. D. Price (2012). Structural Determinants of the Outer Shell of β -Carboxysomes in *Synechococcus elongatus* PCC 7942: Roles for CcmK2, K3-K4, CcmO, and CcmL. *PLoS ONE*, 7(8):e43871
- [24] O. A. Koksharova und C. P. Wolk (2002). Genetic tools for cyanobacteria. *Applied microbiology and biotechnology*, 58(2):123–137
- [25] G. Dong und S. S. Golden (2008). How a cyanobacterium tells time. *Current opinion in microbiology*, 11(6):541–546
- [26] M. Ishiura, S. Kutsuna, S. Aoki, H. Iwasaki, C. R. Andersson et al. (1998). Expression of a gene cluster *kaiABC* as a circadian feedback process in cyanobacteria. *Science (New York, N.Y.)*, 281(5382):1519–1523
- [27] C. K. Holtman, Y. Chen, P. Sandoval, A. Gonzales, M. S. Nalty et al. (2005). High-throughput functional analysis of the *Synechococcus elongatus* PCC 7942 genome. *DNA research: an international journal for rapid publication of reports on genes and genomes*, 12(2):103–115
- [28] C. Brettschneider, R. J. Rose, S. Hertel, I. M. Axmann, A. J. R. Heck et al. (2010). A sequestration feedback determines dynamics and temperature entrainment of the KaiABC circadian clock. *Molecular Systems Biology*, 6:389
- [29] S. Hertel, C. Brettschneider und I. M. Axmann (2013). Revealing a two-loop transcriptional feedback mechanism in the cyanobacterial circadian clock. *PLoS computational biology*, 9(3):e1002966
- [30] R. Pattanayek, D. R. Williams, G. Rossi, S. Weigand, T. Mori et al. (2011). Combined SAXS/EM based models of the *S. elongatus* post-translational circadian oscillator and its interactions with the output His-kinase SasA. *PloS one*, 6(8):e23697
- [31] H. Miyashita, H. Ikemoto, N. Kurano, K. Adachi, M. Chihara et al. (1996). Chlorophyll *d* as a major pigment. *Nature*, 383(6599):402–402

- [32] E. A. Welsh, M. Liberton, J. Stöckel, T. Loh, T. Elvitigala et al. (2008). The genome of *Cyanothece* 51142, a unicellular diazotrophic cyanobacterium important in the marine nitrogen cycle. *Proceedings of the National Academy of Sciences of the United States of America*, 105(39):15094–15099
- [33] S. Schneider und U. J. Jürgens (1991). Cell wall and sheath constituents of the cyanobacterium *Gloeobacter violaceus*. *Archives of Microbiology*, 156(4):312–318
- [34] P. J. Oberholster, A.-M. Botha und J. U. Grobbelaar (2004). *Microcystis aeruginosa* : source of toxic microcystins in drinking water. *African Journal of Biotechnology*, 3(3)
- [35] T. Kaneko, Y. Nakamura, C. P. Wolk, T. Kuritz, S. Sasamoto et al. (2001). Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA research: an international journal for rapid publication of reports on genes and genomes*, 8(5):205–213; 227–253
- [36] F. Partensky, J. Blanchot und D. Vaulot (1999). Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. *Bulletin de l'Institut océanographique, Monaco*, 19:457–475
- [37] J. P. Allewalt, M. M. Bateson, N. P. Revsbech, K. Slack und D. M. Ward (2006). Effect of Temperature and Light on Growth of and Photosynthesis by *Synechococcus* Isolates Typical of Those Predominating in the Octopus Spring Microbial Mat Community of Yellowstone National Park. *Applied and Environmental Microbiology*, 72(1):544–550
- [38] J. Batterton, J C und C. Van Baalen (1971). Growth responses of blue-green algae to sodium chloride concentration. *Archiv für Mikrobiologie*, 76(2):151–165
- [39] J. B. Waterbury, S. W. Watson, R. R. L. Guillard und L. E. Brand (1979). Widespread occurrence of a unicellular, marine, planktonic, cyanobacterium. *Nature*, 277(5694):293–294
- [40] T. Yamaoka, K. Satoh und S. Katoh (1978). Photosynthetic activities of a thermophilic blue-green alga. *Plant and Cell Physiology*, 19(6):943–954
- [41] D. G. Capone, J. P. Zehr, H. W. Paerl, B. Bergman und E. J. Carpenter (1997). *Trichodesmium*, a Globally Significant Marine Cyanobacterium. *Science*, 276(5316):1221–1229
- [42] L. Hood (1998). Systems biology: New opportunities arising from genomics, proteomics and beyond. *Experimental Hematology*, 26(8):681
- [43] H. Kitano, S. Hamahashi und S. Luke (1998). The Perfect *C. elegans* Project: An Initial Report. *Artificial Life*, 4:141–156
- [44] Tomita, Hashimoto, Takahashi, Shimizu, Matsuzaki et al. (1997). E-CELL: Software Environment for Whole Cell Simulation. *Genome informatics. Workshop on Genome Informatics*, 8:147–155
- [45] A. R. Joyce und B. O. Palsson (2006). The model organism as a system: integrating 'omics' data sets. *Nature reviews. Molecular cell biology*, 7(3):198–210
- [46] U. Stelzl und E. E. Wanker (2006). The value of high quality protein-protein interaction networks for systems biology. *Current opinion in chemical biology*, 10(6):551–558
- [47] J. Mestres, E. Gregori-Puigjané, S. Valverde und R. V. Solé (2008). Data completeness—the Achilles heel of drug-target networks. *Nature biotechnology*, 26(9):983–984
- [48] I. Thiele und B. O. Palsson (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protocols*, 5(1):93–121
- [49] D. A. Beard und H. Qian (2004). Constraint-based modeling of metabolomic systems. *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*. John Wiley & Sons, Ltd, 1–10
- [50] I. Schomburg, A. Chang und D. Schomburg (2002). BRENDA, enzyme data and metabolic information. *Nucleic acids research*, 30(1):47–49

- [51] M. Kanehisa und S. Goto (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28(1):27–30
- [52] R. Caspi, T. Altman, J. M. Dale, K. Dreher, C. A. Fulcher et al. (2010). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Research*, 38(suppl 1):D473–D479
- [53] N. C. Duarte, S. A. Becker, N. Jamshidi, I. Thiele, M. L. Mo et al. (2007). Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proceedings of the National Academy of Sciences*, 104(6):1777–1782
- [54] D. A. Fell, M. G. Poolman und A. Gevorgyan (2010). Building and analysing genome-scale metabolic models. *Biochemical Society transactions*, 38(5):1197–1201
- [55] S. F. Altschul, W. Gish, W. Miller, E. W. Myers und D. J. Lipman (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410
- [56] J. D. Thompson, D. G. Higgins und T. J. Gibson (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*, 22(22):4673–4680
- [57] M. Durot, P.-Y. Bourguignon und V. Schachter (2009). Genome-scale models of bacterial metabolism: reconstruction and applications. *FEMS Microbiology Reviews*, 33(1):164–190
- [58] A. M. Feist und B. O. Palsson (2010). The biomass objective function. *Current opinion in microbiology*, 13(3):344–349
- [59] H. Kitano (2000). Perspectives on systems biology. *New Generation Computing*, 18(3):199–216
- [60] T. Shlomi, Y. Eisenberg, R. Sharan und E. Ruppin (2007). A genome-scale computational study of the interplay between transcriptional regulation and metabolism. *Molecular systems biology*, 3:101
- [61] J. S. Edwards, R. U. Ibarra und B. O. Palsson (2001). *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nature biotechnology*, 19(2):125–130
- [62] K. Y. Lee, J. M. Park, T. Y. Kim, H. Yun und S. Y. Lee (2010). The genome-scale metabolic network analysis of *Zymomonas mobilis* ZM4 explains physiological features and suggests ethanol and succinic acid production strategies. *Microbial Cell Factories*, 9(1):94
- [63] O. Wolkenhauer (2001). Systems biology: The reincarnation of systems theory applied in biology? *Briefings in Bioinformatics*, 2(3):258–270
- [64] B. Palsson (2000). The challenges of in silico biology. *Nature biotechnology*, 18(11):1147–1150
- [65] B. L. Clarke (1988). Stoichiometric network analysis. *Cell biophysics*, 12:237–253
- [66] N. D. Price, J. L. Reed und B. O. Palsson (2004). Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nature reviews. Microbiology*, 2(11):886–897
- [67] J. Tollefson (2011). Seven billion and counting. *Nature News*, 478(7369):300–300
- [68] B. Hankamer, F. Lehr, J. Rupprecht, J. H. Mussnug, C. Posten et al. (2007). Photosynthetic biomass and H₂ production by green algae: from bioengineering to bioreactor scale-up. *Physiologia plantarum*, 131(1):10–21
- [69] H. C. Greenwell, L. M. L. Laurens, R. J. Shields, R. W. Lovitt und K. J. Flynn (2010). Placing microalgae on the biofuels priority list: a review of the technological challenges. *Journal of the Royal Society, Interface / the Royal Society*, 7(46):703–726
- [70] N. E. Korres, A. Singh, A.-S. Nizami und J. D. Murphy (2010). Is grass biomethane a sustainable transport biofuel? *Biofuels, Bioproducts and Biorefining*, 4(3):310–325

- [71] C. Posten und G. Schaub (2009). Microalgae and terrestrial biomass as source for fuels—A process view. *Journal of Biotechnology*, 142(1):64–69
- [72] A. Singh, P. S. Nigam und J. D. Murphy (2011). Renewable fuels from algae: An answer to debatable land based fuels. *Bioresource Technology*, 102(1):10–16
- [73] R. Kalscheuer, T. Stölting und A. Steinbüchel (2006). Microdiesel: *Escherichia coli* engineered for fuel production. *Microbiology*, 152(9):2529–2536
- [74] A. Schirmer, M. A. Rude, X. Li, E. Popova und S. B. del Cardayre (2010). Microbial Biosynthesis of Alkanes. *Science*, 329(5991):559–562
- [75] S. Atsumi, A. F. Cann, M. R. Connor, C. R. Shen, K. M. Smith et al. (2008). Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metabolic engineering*, 10(6):305–311
- [76] B. B. Bond-Watts, R. J. Bellerose und M. C. Y. Chang (2011). Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nature chemical biology*, 7(4):222–227
- [77] C. R. Shen, E. I. Lan, Y. Dekishima, A. Baez, K. M. Cho et al. (2011). Driving Forces Enable High-Titer Anaerobic 1-Butanol Synthesis in *Escherichia coli*. *Applied and Environmental Microbiology*, 77(9):2905–2915
- [78] A. Pandey (2009). *Handbook of plant-based biofuels*. CRC Press, Boca Raton
- [79] A. Pandey, C. Larroche, S. C. Ricke, C.-G. Dussap und E. Gnansounou (2011). *Biofuels: Alternative Feedstocks and Conversion Processes*. Academic Press
- [80] J. Hill, E. Nelson, D. Tilman, S. Polasky und D. Tiffany (2006). Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences*, 103(30):11206–11210
- [81] J. P. W. Scharlemann und W. F. Laurance (2008). How Green Are Biofuels? *Science*, 319(5859):43–44
- [82] E. I. Lan und J. C. Liao (2011). Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. *Metabolic Engineering*, 13(4):353–363
- [83] X. Liu, J. Sheng und r. Curtiss, Roy (2011). Fatty acid production in genetically modified cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 108(17):6899–6904
- [84] S. S. Golden, J. Brusslan und R. Haselkorn (1987). Genetic engineering of the cyanobacterial chromosome. *Methods in enzymology*, 153:215–231
- [85] T. Heidorn, D. Camsund, H.-H. Huang, P. Lindberg, P. Oliveira et al. (2011). Synthetic biology in cyanobacteria engineering and analyzing novel functions. *Methods in enzymology*, 497:539–579
- [86] H.-H. Huang, D. Camsund, P. Lindblad und T. Heidorn (2010). Design and characterization of molecular tools for a Synthetic Biology approach towards developing cyanobacterial biotechnology. *Nucleic acids research*, 38(8):2577–2593
- [87] A. M. Ruffing (2011). Engineered cyanobacteria: teaching an old bug new tricks. *Bioengineered bugs*, 2(3):136–149
- [88] A. Parmar, N. K. Singh, A. Pandey, E. Gnansounou und D. Madamwar (2011). Cyanobacteria and microalgae: A positive prospect for biofuels. *Bioresource Technology*, 102(22):10163–10172
- [89] T. Searchinger, R. Heimlich, R. A. Houghton, F. Dong, A. Elobeid et al. (2008). Use of U.S. Croplands for Biofuels Increases Greenhouse Gases Through Emissions from Land-Use Change. *Science*, 319(5867):1238–1240
- [90] P. Tamagnini, E. Leitão, P. Oliveira, D. Ferreira, F. Pinto et al. (2007). Cyanobacterial hydrogenases: diversity, regulation and applications. *FEMS microbiology reviews*, 31(6):692–720

- [91] K. Takahama, M. Matsuoka, K. Nagahama und T. Ogawa (2003). Construction and analysis of a recombinant cyanobacterium expressing a chromosomally inserted gene for an ethylene-forming enzyme at the *psbAI* locus. *Journal of bioscience and bioengineering*, 95(3):302–305
- [92] A. M. Ruffing und H. D. T. Jones (2012). Physiological effects of free fatty acid production in genetically engineered *Synechococcus elongatus* PCC 7942. *Biotechnology and bioengineering*, 109(9):2190–2199
- [93] H. Knoop, Y. Zilliges, W. Lockau und R. Steuer (2010). The Metabolic Network of *Synechocystis* sp. PCC 6803: Systemic Properties of Autotrophic Growth. *Plant Physiol.*, 154(1):410–422
- [94] H. Knoop, M. Gründel, Y. Zilliges, R. Lehmann, S. Hoffmann et al. (2013). Flux Balance Analysis of Cyanobacterial Metabolism: The Metabolic Network of *Synechocystis* sp. PCC 6803. *PLoS Comput Biol*, 9(6):e1003081
- [95] S. Schuster und R. Schuster (1991). Detecting strictly detailed balanced subnetworks in open chemical reaction networks. *Journal of Mathematical Chemistry*, 6(1):17–40
- [96] D. A. Fell und J. R. Small (1986). Fat synthesis in adipose tissue. An examination of stoichiometric constraints. *Biochemical Journal*, 238(3):781–786
- [97] G. Strang (1980). *Linear algebra and its applications*. Academic Press, New York, 2 Aufl.
- [98] C. H. Schilling, D. Letscher und B. O. Palsson (2000). Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *Journal of theoretical biology*, 203(3):229–248
- [99] D. A. Beard, S.-d. Liang und H. Qian (2002). Energy balance for analysis of complex metabolic networks. *Biophysical journal*, 83(1):79–86
- [100] A. Hoppe, S. Hoffmann und H.-G. Holzhütter (2007). Including metabolite concentrations into flux balance analysis: thermodynamic realizability as a constraint on flux distributions in metabolic networks. *BMC Systems Biology*, 1:23
- [101] S. Schuster, T. Pfeiffer und D. A. Fell (2008). Is maximization of molar yield in metabolic networks favoured by evolution? *Journal of theoretical biology*, 252(3):497–504
- [102] A. Raghunathan, J. Reed, S. Shin, B. Palsson und S. Daefler (2009). Constraint-based analysis of metabolic capacity of *Salmonella typhimurium* during host-pathogen interaction. *BMC systems biology*, 3:38
- [103] J. D. Orth, I. Thiele und B. O. Palsson (2010). What is flux balance analysis? *Nat Biotech*, 28(3):245–248
- [104] S. A. Becker, A. M. Feist, M. L. Mo, G. Hannum, B. O. Palsson et al. (2007). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nature Protocols*, 2(3):727–738
- [105] J. Schellenberger, R. Que, R. M. T. Fleming, I. Thiele, J. D. Orth et al. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protocols*, 6(9):1290–1307
- [106] S. Klamt, J. Saez-Rodriguez und E. D. Gilles (2007). Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC Systems Biology*, 1(1):2
- [107] A. Hoppe, S. Hoffmann, A. Gerasch, C. Gille und H.-G. Holzhütter (2011). FASIMU: flexible software for flux-balance computation series in large metabolic networks. *BMC bioinformatics*, 12:28
- [108] I. Gurobi Optimization (2013). Gurobi Optimizer Reference Manual
- [109] D. Segrè, D. Vitkup und G. M. Church (2002). Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(23):15112–15117

- [110] R. Schuetz, L. Kuepfer und U. Sauer (2007). Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Molecular Systems Biology*, 3:119
- [111] A. M. Feist, M. J. Herrgard, I. Thiele, J. L. Reed und B. O. Palsson (2009). Reconstruction of biochemical networks in microorganisms. *Nat Rev Micro*, 7(2):129–143
- [112] M. A. Oberhardt, B. O. Palsson und J. A. Papin (2009). Applications of genome-scale metabolic reconstructions. *Molecular systems biology*, 5:320
- [113] D. A. Benson, I. Karsch-Mizrachi, K. Clark, D. J. Lipman, J. Ostell et al. (2012). GenBank. *Nucleic acids research*, 40(Database issue):D48–53
- [114] Y. Nakamura, T. Kaneko, M. Hirosawa, N. Miyajima und S. Tabata (1998). CyanoBase, a www database containing the complete nucleotide sequence of the genome of *Synechocystis* sp. strain PCC6803. *Nucleic Acids Research*, 26(1):63–67
- [115] M. Nakao, S. Okamoto, M. Kohara, T. Fujishiro, T. Fujisawa et al. (2010). CyanoBase: the cyanobacteria genome database update 2010. *Nucleic Acids Research*, 38(suppl 1):D379–D381
- [116] H. J. Tripp, S. R. Bench, K. A. Turk, R. A. Foster, B. A. Desany et al. (2010). Metabolic streamlining in an open-ocean nitrogen-fixing cyanobacterium. *Nature*, 464(7285):90–94
- [117] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle et al. (2003). The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531
- [118] M. Eisenhut, S. Kahlon, D. Hasse, R. Ewald, J. Lieman-Hurwitz et al. (2006). The Plant-Like C₂ Glycolate Cycle and the Bacterial-Like Glycerate Pathway Cooperate in Phosphoglycolate Metabolism in Cyanobacteria. *Plant Physiology*, 142(1):333–342
- [119] M. Eisenhut, W. Ruth, M. Haimovich, H. Bauwe, A. Kaplan et al. (2008). The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiotically to plants. *Proceedings of the National Academy of Sciences of the United States of America*, 105(44):17199–17204
- [120] A. Shastri und J. Morgan (2005). Flux Balance Analysis of Photoautotrophic Metabolism. *Biotechnology Progress*, 21(6):1617–1626
- [121] S.-J. Hong und C.-G. Lee (2007). Evaluation of central metabolism based on a genomic database of *Synechocystis* PCC 6803. *Biotechnology and Bioprocess Engineering*, 12(2):165–173
- [122] P. Fu (2009). Genome-scale modeling of *Synechocystis* sp. PCC 6803 and prediction of pathway insertion. *Journal of Chemical Technology & Biotechnology*, 84(4):473–483
- [123] A. Montagud, E. Navarro, P. Fernandez de Cordoba, J. Urchueguia und K. Patil (2010). Reconstruction and analysis of genome-scale metabolic model of a photosynthetic bacterium. *BMC Systems Biology*, 4(1):156
- [124] K. Yoshikawa, Y. Kojima, T. Nakajima, C. Furusawa, T. Hirasawa et al. (2011). Reconstruction and verification of a genome-scale metabolic model for *Synechocystis* sp. PCC6803. *Applied microbiology and biotechnology*, 92(2):347–358
- [125] J. Nogales, S. Gudmundsson, E. M. Knight, B. O. Palsson und I. Thiele (2012). Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 109(7):2678–2683
- [126] R. Saha, A. T. Verseput, B. M. Berla, T. J. Mueller, H. B. Pakrasi et al. (2012). Reconstruction and Comparison of the Metabolic Potential of Cyanobacteria *Cyanothece* sp. ATCC 51142 and *Synechocystis* sp. PCC 6803. *PLoS ONE*, 7(10):e48285

- [127] G. Taroncher-Oldenburg, K. Nishina und G. Stephanopoulos (2000). Identification and Analysis of the Polyhydroxyalkanoate-Specific β -Ketothiolase and Acetoacetyl Coenzyme A Reductase Genes in the Cyanobacterium *Synechocystis* sp. Strain PCC6803. *Applied and Environmental Microbiology*, 66(10):4440–4448
- [128] D. Dähnhardt, J. Falk, J. Appel, T. A. van der Kooij, R. Schulz-Friedrich et al. (2002). The hydroxyphenylpyruvate dioxygenase from *Synechocystis* sp. PCC 6803 is not required for plastoquinone biosynthesis. *FEBS Letters*, 523(1-3):177–181
- [129] R. Sadre, C. Pfaff und S. Buchkremer (2012). Plastoquinone-9 biosynthesis in cyanobacteria differs from that in plants and involves a novel 4-hydroxybenzoate solanesyltransferase. *The Biochemical journal*, 442(3):621–629
- [130] C. Beck, H. Knoop, I. M. Axmann und R. Steuer (2012). The diversity of cyanobacterial metabolism: genome analysis of multiple phototrophic microorganisms. *BMC Genomics*, 13(1):56
- [131] Y. Chiba, K. Oshima, H. Arai, M. Ishii und Y. Igarashi (2012). Discovery and analysis of cofactor-dependent phosphoglycerate mutase homologs as novel phosphoserine phosphatases in *Hydrogenobacter thermophilus*. *The Journal of biological chemistry*, 287(15):11934–11941
- [132] N. Swainston, K. Smallbone, P. Mendes, D. Kell und N. Paton (2011). The SuBliMinaL Toolbox: automating steps in the reconstruction of metabolic networks. *Journal of integrative bioinformatics*, 8(2):186
- [133] G. Falkner und F. Horner (1976). pH Changes in the Cytoplasm of the Blue-Green Alga *Anacystis nidulans* Caused by Light-dependent Proton Flux into the Thylakoid Space. *Plant physiology*, 58(6):717–718
- [134] N. L. Novère, A. Finney, M. Hucka, U. S. Bhalla, F. Campagne et al. (2005). Minimum information requested in the annotation of biochemical models (MIRIAM). *Nature Biotechnology*, 23(12):1509–1515
- [135] K. Degtyarenko, P. d. Matos, M. Ennis, J. Hastings, M. Zbinden et al. (2008). ChEBI: a database and ontology for chemical entities of biological interest. *Nucleic Acids Research*, 36(suppl 1):D344–D350
- [136] T. Omata, G. D. Price, M. R. Badger, M. Okamura, S. Gohta et al. (1999). Identification of an ATP-binding cassette transporter involved in bicarbonate uptake in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *Proceedings of the National Academy of Sciences*, 96(23):13571–13576
- [137] H.-G. Holzhütter (2004). The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. *European journal of biochemistry / FEBS*, 271(14):2905–2922
- [138] G. Cogne, J.-B. Gros und C.-G. Dussap (2003). Identification of a metabolic network structure representative of *Arthrospira platensis* metabolism. *Biotechnology and Bioengineering*, 84(6):667–676
- [139] A. Montagud, A. Zelezniak, E. Navarro, P. F. de Córdoba, J. F. Urchueguía et al. (2011). Flux coupling and transcriptional regulation within the metabolic network of the photosynthetic bacterium *Synechocystis* sp. PCC6803. *Biotechnology Journal*, 6(3):330–342
- [140] C. Yang, Q. Hua und K. Shimizu (2002). Metabolic flux analysis in *Synechocystis* using isotope distribution from ^{13}C -labeled glucose. *Metabolic Engineering*, 4(3):202–216
- [141] J. Y. Yen, H. Nazem-Bokaei, B. G. Freedman, A. I. M. Athamneh und R. S. Senger (2013). Deriving Metabolic Engineering Strategies from Genome-Scale Modeling with Flux Ratio Constraints. *Biotechnology journal*
- [142] S. Zhang und D. A. Bryant (2011). The Tricarboxylic Acid Cycle in Cyanobacteria. *Science*, 334(6062):1551–1553

- [143] T. T. Vu, S. M. Stolyar, G. E. Pinchuk, E. A. Hill, L. A. Kucek et al. (2012). Genome-Scale Modeling of Light-Driven Reductant Partitioning and Carbon Fluxes in Diazotrophic Unicellular Cyanobacterium *Cyanothece* sp. ATCC 51142. *PLoS Comput Biol*, 8(4):e1002460
- [144] J. J. Hamilton und J. L. Reed (2012). Identification of Functional Differences in Metabolic Networks Using Comparative Genomics and Constraint-Based Models. *PLoS ONE*, 7(4):e34670
- [145] R. Steuer, H. Knoop und R. Machné (2012). Modelling cyanobacteria: from metabolism to integrative models of phototrophic growth. *Journal of experimental botany*, 63(6):2259–2274
- [146] J. W. Cooley und W. F. J. Vermaas (2001). Succinate Dehydrogenase and Other Respiratory Pathways in Thylakoid Membranes of *Synechocystis* sp. Strain PCC 6803: Capacity Comparisons and Physiological Function. *J. Bacteriol.*, 183(14):4251–4258
- [147] W. F. Vermaas (2001). Photosynthesis and Respiration in Cyanobacteria. *eLS*. John Wiley & Sons, Ltd
- [148] D. C. Yao, D. C. Brune und W. F. Vermaas (2012). Lifetimes of photosystem I and II proteins in the cyanobacterium *Synechocystis* sp. PCC 6803. *FEBS Letters*, 586(2):169–173
- [149] Y. Helman, D. Tchernov, L. Reinhold, M. Shibata, T. Ogawa et al. (2003). Genes encoding A-type flavoproteins are essential for photoreduction of O₂ in cyanobacteria. *Current biology: CB*, 13(3):230–235
- [150] Y. Allahverdiyeva, M. Ermakova, M. Eisenhut, P. Zhang, P. Richaud et al. (2011). Interplay between Flavodiiron Proteins and Photorespiration in *Synechocystis* sp. PCC 6803. *The Journal of Biological Chemistry*, 286(27):24007–24014
- [151] D. Tchernov, J. Silverman, B. Luz, L. Reinhold und A. Kaplan (2003). Massive light-dependent cycling of inorganic carbon between oxygenic photosynthetic microorganisms and their surroundings. *Photosynthesis research*, 77(2-3):95–103
- [152] Y. Helman, E. Barkan, D. Eisenstadt, B. Luz und A. Kaplan (2005). Fractionation of the three stable oxygen isotopes by oxygen-producing and oxygen-consuming reactions in photosynthetic organisms. *Plant physiology*, 138(4):2292–2298
- [153] M. G. Poolman, D. A. Fell und C. A. Raines (2003). Elementary modes analysis of photosynthate metabolism in the chloroplast stroma. *European Journal of Biochemistry*, 270(3):430–439
- [154] O. Koksharova, M. Schubert, S. Shestakov und R. Cerff (1998). Genetic and biochemical evidence for distinct key functions of two highly divergent GAPDH genes in catabolic and anabolic carbon flow of the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Molecular Biology*, 36(1):183–194
- [155] L. Van Liere und A. Walsby (1982). Interactions of cyanobacteria with light. *Carr NG, Whitton BA, The Biology of Cyanobacteria*, Bd. 19 von *Botanical Monographs*. Blackwell Scientific Publications, Oxford, 2 Aufl., 9–46
- [156] J. D. Young, A. A. Shastri, G. Stephanopoulos und J. A. Morgan (2011). Mapping photoautotrophic metabolism with isotopically nonstationary ¹³C flux analysis. *Metabolic engineering*, 13(6):656–665
- [157] M. Sonderegger, M. Schümperli und U. Sauer (2004). Metabolic engineering of a phosphoketolase pathway for pentose catabolism in *Saccharomyces cerevisiae*. *Applied and environmental microbiology*, 70(5):2892–2897
- [158] A. Chinen, Y. I. Kozlov, Y. Hara, H. Izui und H. Yasueda (2007). Innovative metabolic pathway design for efficient l-glutamate production by suppressing CO₂ emission. *Journal of bioscience and bioengineering*, 103(3):262–269

- [159] E. Fischer und U. Sauer (2005). Large-scale in vivo flux analysis shows rigidity and suboptimal performance of *Bacillus subtilis* metabolism. *Nature genetics*, 37(6):636–640
- [160] C. Hackenberg, A. Engelhardt, H. C. P. Matthijs, F. Wittink, H. Bauwe et al. (2009). Photorespiratory 2-phosphoglycolate metabolism and photoreduction of O₂ cooperate in high-light acclimation of *Synechocystis* sp. strain PCC 6803. *Planta*, 230(4):625–637
- [161] H. Bauwe, M. Hagemann, R. Kern und S. Timm (2012). Photorespiration has a dual origin and manifold links to central metabolism. *Current opinion in plant biology*, 15(3):269–275
- [162] M. R. Badger und G. D. Price (2003). CO₂ concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *Journal of Experimental Botany*, 54(383):609–622
- [163] J. Huege, J. Goetze, D. Schwarz, H. Bauwe, M. Hagemann et al. (2011). Modulation of the Major Paths of Carbon in Photorespiratory Mutants of *Synechocystis*. *PLoS ONE*, 6(1):e16278
- [164] B. Colman und E. G. Norman (1997). Serine synthesis in cyanobacteria by a nonphotorespiratory pathway. *Physiologia Plantarum*, 100(1):133–136
- [165] M. Gründel, R. Scheunemann, W. Lockau und Y. Zilliges (2012). Impaired glycogen synthesis causes metabolic overflow reactions and affects stress responses in the cyanobacterium *Synechocystis* sp. PCC 6803. *Microbiology (Reading, England)*, 158(Pt 12):3032–3043
- [166] J. Pearce und N. G. Carr (1967). The Metabolism of Acetate by the Blue-green Algae, *Anabaena variabilis* and *Anacystis nidulans*. *J Gen Microbiol*, 49(2):301–313
- [167] J. H. Eley (1988). Glyoxylate Cycle Enzyme Activities in the Cyanobacterium *Anacystis Nidulans*. *Journal of Phycology*, 24(4):586–588
- [168] A. Bandyopadhyay, T. Elvitigala, E. Welsh, J. Stöckel, M. Liberton et al. (2011). Novel metabolic attributes of the genus *Cyanothece*, comprising a group of unicellular nitrogen-fixing Cyanobacteria. *mBio*, 2(5):1–10
- [169] G. Dixon und H. Kornberg (1959). Assay methods for key enzymes of the glyoxylate cycle. *Biochem. J*, 72(3):151–159
- [170] L. J. Sweetlove, K. F. Beard, A. Nunes-Nesi, A. R. Fernie und R. G. Ratcliffe (2010). Not just a circle: flux modes in the plant TCA cycle. *Trends in Plant Science*, 15(8):462–470
- [171] G. E. Murphy und G. J. Jensen (2005). Electron Cryotomography of the *E. coli* Pyruvate and 2-Oxoglutarate Dehydrogenase Complexes. *Structure*, 13(12):1765–1773
- [172] T. S. Najdi, G. W. Hatfield und E. D. Mjolsness (2010). A 'random steady-state' model for the pyruvate dehydrogenase and alpha-ketoglutarate dehydrogenase enzyme complexes. *Physical biology*, 7:16016
- [173] S. Schuster, L. F. de Figueiredo, A. Schroeter und C. Kaleta (2011). Combining metabolic pathway analysis with Evolutionary Game Theory: explaining the occurrence of low-yield pathways by an analytic optimization approach. *Bio Systems*, 105(2):147–153
- [174] C. H. Johnson, P. L. Stewart und M. Egli (2011). The cyanobacterial circadian system: from biophysics to bioevolution. *Annual review of biophysics*, 40:143–167
- [175] J. Stöckel, E. A. Welsh, M. Liberton, R. Kunnvakkam, R. Aurora et al. (2008). Global transcriptomic analysis of *Cyanothece* 51142 reveals robust diurnal oscillation of central metabolic processes. *Proceedings of the National Academy of Sciences*, 105(16):6156–6161
- [176] J. Cervený und L. Nedbal (2009). Metabolic rhythms of the cyanobacterium *Cyanothece* sp. ATCC 51142 correlate with modeled dynamics of circadian clock. *Journal of Biological Rhythms*, 24(4):295–303

- [177] J. Stöckel, J. M. Jacobs, T. R. Elvitigala, M. Liberton, E. A. Welsh et al. (2011). Diurnal Rhythms Result in Significant Changes in the Cellular Protein Complement in the Cyanobacterium *Cyanothece* 51142. *PLoS ONE*, 6(2):e16680
- [178] J. E. McDermott, C. S. Oehmen, L. A. McCue, E. Hill, D. M. Choi et al. (2011). A model of cyclic transcriptomic behavior in the cyanobacterium *Cyanothece* sp. ATCC 51142. *Molecular bioSystems*, 7(8):2407–2418
- [179] J. R. Waldbauer, S. Rodrigue, M. L. Coleman und S. W. Chisholm (2012). Transcriptome and proteome dynamics of a light-dark synchronized bacterial cell cycle. *PloS one*, 7(8):e43432
- [180] A. S. Blazier und J. A. Papin (2012). Integration of expression data in genome-scale metabolic network reconstructions. *Frontiers in physiology*, 3:299
- [181] R. Lehmann, R. Machné, J. Georg, M. Benary, I. Axmann et al. (2013). How cyanobacteria pose new problems to old methods: challenges in microarray time series analysis. *BMC bioinformatics*, 14:133
- [182] T. Mori, B. Binder und C. H. Johnson (1996). Circadian gating of cell division in cyanobacteria growing with average doubling times of less than 24 hours. *Proceedings of the National Academy of Sciences of the United States of America*, 93(19):10183–10188
- [183] R. Mahadevan, J. S. Edwards und r. Doyle, Francis J (2002). Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophysical journal*, 83(3):1331–1340
- [184] M. Stitt, R. Sulpice und J. Keurentjes (2010). Metabolic Networks: How to Identify Key Components in the Regulation of Metabolism and Growth. *Plant Physiol.*, 152(2):428–444
- [185] K. Nakahigashi, Y. Toya, N. Ishii, T. Soga, M. Hasegawa et al. (2009). Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism. *Molecular systems biology*, 5:306
- [186] D. A. Benson, M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, D. J. Lipman et al. (2013). GenBank. *Nucleic acids research*, 41(Database issue):D36–42
- [187] T. Kaneko, Y. Nakamura, S. Sasamoto, A. Watanabe, M. Kohara et al. (2003). Structural analysis of four large plasmids harboring in a unicellular cyanobacterium, *Synechocystis* sp. PCC 6803. *DNA research: an international journal for rapid publication of reports on genes and genomes*, 10(5):221–228
- [188] R. S. Gupta und D. W. Mathews (2010). Signature proteins for the major clades of Cyanobacteria. *BMC evolutionary biology*, 10:24
- [189] V. M. Markowitz, I.-M. A. Chen, K. Palaniappan, K. Chu, E. Szeto et al. (2010). The integrated microbial genomes system: an expanding comparative analysis resource. *Nucleic acids research*, 38(Database issue):D382–390
- [190] R. Rippka, J. Deruelles, J. B. Waterbury, M. Herdman und R. Y. Stanier (1979). Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *J Gen Microbiol*, 111(1):1–61
- [191] Y. Moriya, M. Itoh, S. Okuda, A. C. Yoshizawa und M. Kanehisa (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, 35(Web Server):W182–W185
- [192] R. R. Sokal, C. D. Michener und U. o. Kansas (1958). *A Statistical Method for Evaluating Systematic Relationships*. University of Kansas
- [193] F. Baumdicker, W. R. Hess und P. Pfaffelhuber (2010). The diversity of a distributed genome in bacterial populations. *The Annals of Applied Probability*, 20(5):1567–1606
- [194] M. L. Coleman, M. B. Sullivan, A. C. Martiny, C. Steglich, K. Barry et al. (2006). Genomic islands and the ecology and evolution of *Prochlorococcus*. *Science (New York, N.Y.)*, 311(5768):1768–1770

- [195] J. S. Hogg, F. Z. Hu, B. Janto, R. Boissy, J. Hayes et al. (2007). Characterization and modeling of the *Haemophilus influenzae* core and supragenomes based on the complete genomic sequences of Rd and 12 clinical nontypeable strains. *Genome biology*, 8(6):R103
- [196] A. Kislyuk, B. Haegeman, N. Bergman und J. Weitz (2011). Genomic fluidity: an integrative view of gene diversity within microbial populations. *BMC Genomics*, 12(1):32
- [197] J. Larsson, J. A. Nylander und B. Bergman (2011). Genome fluctuations in cyanobacteria reflect evolutionary, developmental and adaptive traits. *BMC Evolutionary Biology*, 11:187
- [198] A. Y. Mulikidjanian, E. V. Koonin, K. S. Makarova, S. L. Mekhedov, A. Sorokin et al. (2006). The cyanobacterial genome core and the origin of photosynthesis. *Proceedings of the National Academy of Sciences*, 103(35):13126–13131
- [199] J. Raymond, O. Zhaxybayeva, J. P. Gogarten, S. Y. Gerdes und R. E. Blankenship (2002). Whole-genome analysis of photosynthetic prokaryotes. *Science (New York, N.Y.)*, 298(5598):1616–1620
- [200] R. L. Tatusov, E. V. Koonin und D. J. Lipman (1997). A genomic perspective on protein families. *Science (New York, N.Y.)*, 278(5338):631–637
- [201] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics*, 25(1):25–29
- [202] M. A. Harris, J. Clark, A. Ireland, J. Lomax, M. Ashburner et al. (2004). The Gene Ontology (GO) database and informatics resource. *Nucleic acids research*, 32(Database issue):D258–261
- [203] A. Alexa, J. Rahnenführer und T. Lengauer (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics (Oxford, England)*, 22(13):1600–1607
- [204] S. Grossmann, S. Bauer, P. N. Robinson und M. Vingron (2007). Improved detection of overrepresentation of Gene-Ontology annotations with parent child analysis. *Bioinformatics (Oxford, England)*, 23(22):3024–3031
- [205] J. Felsenstein (2005). PHYLIP (Phylogeny Inference Package) version 3.6
- [206] J. R. Cole, Q. Wang, E. Cardenas, J. Fish, B. Chai et al. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic acids research*, 37(Database issue):D141–145
- [207] R. L. Tatusov, D. A. Natale, I. V. Garkavtsev, T. A. Tatusova, U. T. Shankavaram et al. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic acids research*, 29(1):22–28
- [208] P. Lapierre und J. P. Gogarten (2009). Estimating the size of the bacterial pan-genome. *Trends in genetics: TIG*, 25(3):107–110
- [209] H. Tettelin, D. Riley, C. Cattuto und D. Medini (2008). Comparative genomics: the bacterial pan-genome. *Current opinion in microbiology*, 11(5):472–477
- [210] H. Tettelin, V. Masignani, M. J. Cieslewicz, C. Donati, D. Medini et al. (2005). Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial “pan-genome”. *Proceedings of the National Academy of Sciences of the United States of America*, 102(39):13950–13955
- [211] A. Mira, A. B. Martín-Cuadrado, G. D’Auria und F. Rodríguez-Valera (2010). The bacterial pan-genome: a new paradigm in microbiology. *International microbiology: the official journal of the Spanish Society for Microbiology*, 13(2):45–57
- [212] X. Chi, Q. Yang, F. Zhao, S. Qin, Y. Yang et al. (2008). Comparative Analysis of Fatty Acid Desaturases in Cyanobacterial Genomes. *Comparative and Functional Genomics*, 2008:1–25

- [213] O. Zhaxybayeva, J. P. Gogarten, R. L. Charlebois, W. F. Doolittle und R. T. Papke (2006). Phylogenetic analyses of cyanobacterial genomes: quantification of horizontal gene transfer events. *Genome research*, 16(9):1099–1108
- [214] T. Shi und P. G. Falkowski (2008). Genome evolution in cyanobacteria: the stable core and the variable shell. *Proceedings of the National Academy of Sciences of the United States of America*, 105(7):2510–2515
- [215] Z. I. Johnson und Y. Lin (2009). *Prochlorococcus*: Approved for export. *Proceedings of the National Academy of Sciences*, 106(26):10400–10401
- [216] A. C. Martiny, S. Kathuria und P. M. Berube (2009). Widespread metabolic potential for nitrite and nitrate assimilation among *Prochlorococcus* ecotypes. *Proceedings of the National Academy of Sciences*, 106(26):10787–10792
- [217] M. Miyake, K. Takase, M. Narato, E. Khatipov, J. Schnackenberg et al. (2000). Polyhydroxybutyrate production from carbon dioxide by cyanobacteria. *Applied Biochemistry and Biotechnology*, 84-86(1):991–1002
- [218] M. I. Muro-Pastor, J. C. Reyes und F. J. Florencio (2005). Ammonium assimilation in cyanobacteria. *Photosynthesis research*, 83(2):135–150
- [219] S. Takaichi und M. Mochimaru (2007). Carotenoids and carotenogenesis in cyanobacteria: unique ketocarotenoids and carotenoid glycosides. *Cellular and Molecular Life Sciences: CMLS*, 64(19-20):2607–2619
- [220] H. Ito, M. Mutsuda, Y. Murayama, J. Tomita, N. Hosokawa et al. (2009). Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in *Synechococcus elongatus*. *Proceedings of the National Academy of Sciences*, 106(33):14168–14173
- [221] Y. Liu, N. F. Tsinoremas, C. H. Johnson, N. V. Lebedeva, S. S. Golden et al. (1995). Circadian orchestration of gene expression in cyanobacteria. *Genes & development*, 9(12):1469–1478
- [222] Y. Nakahira, M. Katayama, H. Miyashita, S. Kutsuna, H. Iwasaki et al. (2004). Global gene repression by KaiC as a master process of prokaryotic circadian system. *Proceedings of the National Academy of Sciences of the United States of America*, 101(3):881–885
- [223] S. E. Sattler, E. B. Cahoon, S. J. Coughlan und D. DellaPenna (2003). Characterization of tocopherol cyclases from higher plants and cyanobacteria. Evolutionary implications for tocopherol synthesis and function. *Plant physiology*, 132(4):2184–2195
- [224] B. Wu, B. Zhang, X. Feng, J. R. Rubens, R. Huang et al. (2010). Alternative isoleucine synthesis pathway in cyanobacterial species. *Microbiology*, 156(2):596–602
- [225] A. L. Guerreiro, M. Benevento, B. van Breukelen, A. M. Altelaar, A. J. Heck et al. (2013). Circadian rhythm in the cyanobacteria *S. elongatus* probed by high-resolution mass spectrometry based proteomics reveals a small-defined set of circadian proteins. *Vorabdruck*
- [226] R. Vidal, L. López-Maury, M. G. Guerrero und F. J. Florencio (2009). Characterization of an alcohol dehydrogenase from the Cyanobacterium *Synechocystis* sp. strain PCC 6803 that responds to environmental stress conditions via the Hik34-Rre1 two-component system. *Journal of bacteriology*, 191(13):4383–4391
- [227] D. M. Warui, N. Li, H. Nørgaard, C. Krebs, J. Bollinger, J. Martin et al. (2011). Detection of formate, rather than carbon monoxide, as the stoichiometric coproduct in conversion of fatty aldehydes to alkanes by a cyanobacterial aldehyde decarbonylase. *Journal of the American Chemical Society*, 133(10):3316–3319
- [228] D. Burgess (retrieved 2013). Thermochemical Data. *NIST Chemistry WebBook*, Hg. P. L. und W.G. Mallard, Nr. 69 in NIST Standard Reference Database. National Institute of Standards and Technology, Gaithersburg MD, 20899

- [229] N. E. Lewis, H. Nagarajan und B. O. Palsson (2012). Constraining the metabolic genotype-phenotype relationship using a phylogeny of *in silico* methods. *Nature reviews. Microbiology*, 10(4):291–305
- [230] J. Kämäräinen, H. Knoop, N. J. Stanford, F. Guerrero, M. K. Akhtar et al. (2012). Physiological tolerance and stoichiometric potential of cyanobacteria for hydrocarbon fuel production. *Journal of biotechnology*, 162(1):67–74
- [231] K. Awai, T. Kakimoto, C. Awai, T. Kaneko, Y. Nakamura et al. (2006). Comparative Genomic Analysis Revealed a Gene for Monoglucosyldiacylglycerol Synthase, an Enzyme for Photosynthetic Membrane Lipid Synthesis in Cyanobacteria. *Plant Physiology*, 141(3):1120–1127
- [232] S. Jantaro, P. Mäenpää, P. Mulo und A. Incharoensakdi (2003). Content and biosynthesis of polyamines in salt and osmotically stressed cells of *Synechocystis* sp. PCC 6803. *FEMS Microbiology Letters*, 228(1):129–135
- [233] G. Mittenhuber (2001). Phylogenetic analyses and comparative genomics of vitamin B₆ (pyridoxine) and pyridoxal phosphate biosynthesis pathways. *Journal of Molecular Microbiology and Biotechnology*, 3(1):1–20
- [234] D. A. Rodionov, A. G. Vitreschak, A. A. Mironov und M. S. Gelfand (2003). Comparative Genomics of the Vitamin B₁₂ Metabolism and Regulation in Prokaryotes. *Journal of Biological Chemistry*, 278(42):41148–41159
- [235] S. Schriek, C. Rückert, D. Staiger, E. K. Pistorius und K.-P. Michel (2007). Bioinformatic evaluation of L-arginine catabolic pathways in 24 cyanobacteria and transcriptional analysis of genes encoding enzymes of L-arginine catabolism in the cyanobacterium *Synechocystis* sp. PCC 6803. *BMC Genomics*, 8:437
- [236] S. Schriek, U. Kahmann, D. Staiger, E. K. Pistorius und K.-P. Michel (2009). Detection of an L-amino acid dehydrogenase activity in *Synechocystis* sp. PCC 6803. *J. Exp. Bot.*, 60(3):1035–1046
- [237] H. Ito, M. Yokono, R. Tanaka und A. Tanaka (2008). Identification of a Novel Vinyl Reductase Gene Essential for the Biosynthesis of Monovinyl Chlorophyll in *Synechocystis* sp. PCC6803. *Journal of Biological Chemistry*, 283(14):9002–9011
- [238] M. R. Islam, S. Aikawa, T. Midorikawa, Y. Kashino, K. Satoh et al. (2008). slr1923 of *Synechocystis* sp. PCC6803 Is Essential for Conversion of 3,8-Divinyl(proto)chlorophyll(ide) to 3-Monovinyl(proto)chlorophyll(ide). *PLANT PHYSIOLOGY*, 148(2):1068–1081
- [239] I. Levin, M. Mevarech und B. A. Palfey (2007). Characterization of a novel bifunctional dihydropteroate synthase/dihydropteroate reductase enzyme from *Helicobacter pylori*. *Journal of Bacteriology*, 189(11):4062–4069
- [240] B. Marteyn, F. Domain, P. Legrain, F. Chauvat und C. Cassier-Chauvat (2009). The thio-redoxin reductase-glutaredoxins-ferredoxin crossroad pathway for selenate tolerance in *Synechocystis* PCC6803. *Molecular Microbiology*, 71(2):520–532
- [241] Y. Nakamura, M. Tsuchiya und H. Ohta (2007). Plastidic Phosphatidic Acid Phosphatases Identified in a Distinct Subfamily of Lipid Phosphate Phosphatases with Prokaryotic Origin. *Journal of Biological Chemistry*, 282(39):29013–29021
- [242] S. Güler, B. Essigmann und C. Benning (2000). A Cyanobacterial Gene, *sqdX*, Required for Biosynthesis of the Sulfolipid Sulfoquinovosyldiacylglycerol. *Journal of Bacteriology*, 182(2):543–545
- [243] J. E. Lunn (2002). Evolution of Sucrose Synthesis. *Plant Physiology*, 128(4):1490–1500
- [244] A. Chatterjee, M. Majee, S. Ghosh und A. L. Majumder (2004). sll1722, an unassigned open reading frame of *Synechocystis* PCC 6803, codes for L-myoinositol 1-phosphate synthase. *Planta*, 218(6):989–998

- [245] A. Chatterjee, K. G. Dastidar, S. Maitra, A. Das-Chatterjee, H. Dihazi et al. (2006). sll1981, an acetolactate synthase homologue of *Synechocystis* sp. PCC6803, functions as L-*myo*-inositol 1-phosphate synthase. *Planta*, 224(2):367–379
- [246] B. Patra, K. Ghosh Dastidar, S. Maitra, J. Bhattacharyya und A. L. Majumder (2007). Functional identification of sll1383 from *Synechocystis* sp PCC 6803 as L-*myo*-inositol 1-phosphate phosphatase (EC 3.1.3.25): molecular cloning, expression and characterization. *Planta*, 225(6):1547–1558
- [247] P. Roughan (1985). Cytidine triphosphate-dependent, acyl-CoA-independent synthesis of phosphatidylglycerol by chloroplasts isolated from spinach and pea. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism*, 835(3):527–532
- [248] J. Browse, N. Warwick, C. R. Somerville und C. R. Slack (1986). Fluxes through the prokaryotic and eukaryotic pathways of lipid synthesis in the '16:3' plant *Arabidopsis thaliana*. *The Biochemical Journal*, 235(1):25–31
- [249] K. Awai, H. Watanabe, C. Benning und I. Nishida (2007). Digalactosyldiacylglycerol is Required for Better Photosynthetic Growth of *Synechocystis* sp. PCC6803 Under Phosphate Limitation. *Plant and Cell Physiology*, 48(11):1517–1523
- [250] M. Merkamm, C. Chassagnole, N. D. Lindley und A. Guyonvarch (2003). Ketopantoate reductase activity is only encoded by *ilvC* in *Corynebacterium glutamicum*. *Journal of Biotechnology*, 104(1–3):253–260
- [251] A. E. Marolewski, K. M. Mattia, M. S. Warren und S. J. Benkovic (1997). Formyl Phosphate: A Proposed Intermediate in the Reaction Catalyzed by *Escherichia coli* PurT GAR Transformylase†. *Biochemistry*, 36(22):6709–6716
- [252] K. Kato, R. Tanaka, S. Sano, A. Tanaka und H. Hosaka (2010). Identification of a gene essential for protoporphyrinogen IX oxidase activity in the cyanobacterium *Synechocystis* sp. PCC6803. *Proceedings of the National Academy of Sciences of the United States of America*, 107(38):16649–16654
- [253] W. Raksajit, P. Yodsang, P. Maenpaa und A. Incharoensakdi (2009). Characterization of spermidine transport system in a cyanobacterium, *Synechocystis* sp. PCC 6803. *Journal of Microbiology and Biotechnology*, 19(5):447–454
- [254] A.-M. Brandt, W. Raksajit, P. Yodsang, P. Mulo, A. Incharoensakdi et al. (2010). Characterization of the substrate-binding PotD subunit in *Synechocystis* sp. strain PCC 6803. *Archives of Microbiology*, 192(10):791–801
- [255] H. Mi, T. Endo, U. Schreiber, T. Ogawa und K. Asada (1992). Electron Donation from Cyclic and Respiratory Flows to the Photosynthetic Intersystem Chain is Mediated by Pyridine Nucleotide Dehydrogenase in the Cyanobacterium *Synechocystis* PCC 6803. *Plant and Cell Physiology*, 33(8):1233–1237
- [256] T. Ogawa und H. Mi (2007). Cyanobacterial NADPH dehydrogenase complexes. *Photosynthesis Research*, 93:69–77
- [257] N. Battchikova, L. Wei, L. Du, L. Bersanini, E.-M. Aro et al. (2011). Identification of Novel Ssl0352 Protein (NdhS), Essential for Efficient Operation of Cyclic Electron Transport around Photosystem I, in NADPH:plastoquinone Oxidoreductase (NDH-1) Complexes of *Synechocystis* sp. PCC 6803. *Journal of Biological Chemistry*, 286(42):36992–37001
- [258] A. D. Lawrence, E. Deery, K. J. McLean, A. W. Munro, R. W. Pickersgill et al. (2008). Identification, Characterization, and Structure/Function Analysis of a Corrin Reductase Involved in Adenosylcobalamin Biosynthesis. *Journal of Biological Chemistry*, 283(16):10813–10821

Danksagung

An dieser Stelle möchte ich mich herzlich bei meinem Doktorvater Prof. Dr. Wolfgang Lockau und bei Prof. Dr. Hanspeter Herzel bedanken, für die Möglichkeit, am Institut für Theoretische Biologie meine Dissertation anfertigen zu dürfen und die Bereitschaft, meiner Promotionskommission beizuwohnen.

Mein ganz besonderer Dank geht an Dr. Ralf Steuer, für seine allumfassende, unermüdliche Hilfe, für die allzeitige Unterstützung und die kompetente wissenschaftliche Betreuung. Vielen Dank für all die Ratschläge und die vielen Anregungen während meiner Arbeit!

Ein großer Dank geht auch an Sabrina Hoffmann, Robert Lehmann, Neil Swainston und besonders Christian Beck für viele fruchtbare Diskussionen und die gute Zusammenarbeit.

Besonderer Dank geht an Dr. Yvonne Zilliges und Dr. Marianne Gründel und den Rest der Arbeitsgruppe 'Biochemie der Pflanzen' für die Durchführung der Experimente und Analysen, die im Laufe meiner Arbeit im Labor anfielen sowie die kompetente Beratung.

Auch herzlichen Dank an Prof. Dr. Martin Hagemann und Dr. Marion Eisenhut für den fachlichen Beistand und die vielen Antworten zu meinen Fragen rund um die Cyanobakterien und speziell zur Photorespiration und den Phosphoserin-Weg.

Den ehemaligen Mitglieder des FORSYS-Partner-Projekts: Systembiologie der Produktion von Biokraftstoffen durch Cyanobakterien und den Mitgliedern des DirectFuel-Projekts, besonders Dr. Patrik Jones gilt mein Dank für die Unterstützung und fachlichen Diskussionen vor allem bei der Netzwerkrekonstruktion und der Analyse der Biokraftstoffsynthesen.

Natürlich möchte ich mich auch bei der DFG und der Europäischen Kommission bedanken, die es mir überhaupt ermöglichten, ohne finanzielle Not an diesem Projekt zu arbeiten und nebenbei die verschiedensten Orte der Welt im Rahmen des wissenschaftlichen Austausches zu besuchen und dabei neue Leute kennenzulernen.

Was nützt die interessanteste Arbeit ohne die freundschaftliche Atmosphäre am Arbeitsplatz? Deswegen will ich mich ganz herzlich bei meinen Bürokollegen und der Arbeitsgruppe 'Metabolic Network Analysis' bedanken, ganz besonders Ilka, Stefanie, Nils und erneut Christian für die schöne Zeit auch außerhalb des ITBs, wie zum Beispiel bei hart umkämpften Matches auf dem Badmintonplatz.

Ein besonders lieber Dank geht an Jasmin Weisemann und Alexander Ahlert, die mich tatkräftig bei dieser Arbeit unterstützten und immer wieder neu motivierten. Vielen Dank für Eure Unterstützung und für die schöne Zeit neben der Arbeit.

Zum Abschluss gilt mein aller größter Dank meinen Eltern, die mich bis zum heutigen Zeitpunkt in jeder Hinsicht unterstützt haben, an mich glaubten und mir alles ermöglicht haben. Ich danke Euch von ganzem Herzen!

Selbständigkeitserklärung

Hiermit versichere ich, die vorliegende Arbeit selbständig angefertigt und keine weiteren als die angegebenen Hilfsmittel und Quellen verwendet zu haben.

Berlin, 9. Februar 2014

.....

Henning Knoop